RESEARCH ARTICLE

Reconstruction of Partially Broken Vascular Structures in X-Ray Images via Vesselness-Loss-Based Multi-Scale Generative Adversarial Networks

KYUNGHOON HAN[1,2], **(Member, IEEE), HEEJOON KOO**[2], **SUNGHEE JUNG**[3], **(Member, IEEE),**
HYUNG-BOK PARK[2,4], **YOUNGTAEK HONG**[2,5], **(Member, IEEE), HACKJOON SHIM**[2],
BYUNGHWAN JEON[6], **(Member, IEEE), AND HYUK-JAE CHANG**[2,5,7]

[1]Department of Internal Medicine, Graduate School of Medical Science, Brain Korea 21 Project, Yonsei University College of Medicine, Seoul 03722, South Korea
[2]CONNECT-AI Research Center, Yonsei University College of Medicine, Seoul 03721, South Korea
[3]School of Electrical Engineering, Kookmin University, Seoul 02707, South Korea
[4]Department of Cardiology, Catholic Kwandong University International St. Mary's Hospital, Incheon 22711, South Korea
[5]Ontact Health Inc., Seoul, South Korea
[6]Division of Computer Engineering, Hankuk University of Foreign Studies, Yongin-si 17035, South Korea
[7]Division of Cardiology, Severance Cardiovascular Hospital, Yonsei University College of Medicine, Yonsei University Health System, Seoul 03721, South Korea

Corresponding author: Byunghwan Jeon (bhjeon@hufs.ac.kr)

ABSTRACT Coronary artery procedures are primarily performed based on X-ray angiography images. However, coronary arteries in X-ray images are often partially broken, complicating diagnoses and procedures owing to lack of visibility. In this paper, we propose a fully automatic method to restore locally broken parts of coronary arteries in X-ray images without using any external information, such as computed tomography images. To this end, we design a new multi-scale generative adversarial network and a vesselness-loss function. The proposed method is optimized for focus on elongated structures and can be utilized in various clinical applications. The proposed method is evaluated and compared with four other existing methods using the performance metrics, PSNR, MSE, and SSIM, and the result shows 34.3, 0.18, and 0.91 averages, respectively for each metric. Based on the performance result, the blocked regions are plausibly reconstructed into such original shapes of blood vessels, which can aid in image-based guiding catheter manipulation during coronary artery procedures. Eventually, the proposed method can be utilized in various clinical applications, e.g., image-based planning and guidance of coronary procedures and prior simulation of results.

INDEX TERMS Coronary artery, X-ray angiography, procedure guidance.

I. INTRODUCTION

Coronary artery diseases (CADs), which are one of the main causes of deaths worldwide [1], [2], [3], result from the blockage of blood vessels by plaques, e.g., cholesterol or fat. Fortunately, such blockages can be removed via

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

percutaneous coronary intervention (PCI) in conjunction with accurate diagnosis. Currently, the two-dimensional (2D) X-ray modality is regarded as the gold standard for PCI image guidance.

2D X-ray angiography is used to identify positions of catheters and patient-specific structures of coronary arteries in real time for PCI guidance. However, it involves injecting contrastive substances, which creates significant burden on
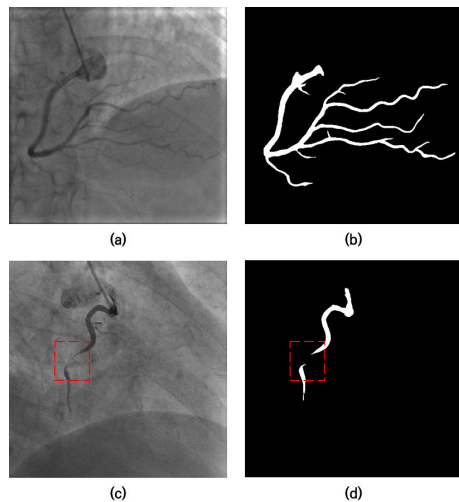
FIGURE 1. Examples of two X-ray images and their segmentation results, depicting broken parts: (a) X-ray image depicting coronary artery without any broken regions; (b) the corresponding segmentation result to (a); (c) An X-ray image depicting the coronary artery with broken parts due to chronic total occlusion (CTO) (red box); (d) the corresponding segmentation result to (c) (red box).

patients, and long-term exposure to X-ray radiation, which is also known to be harmful to both patients and operators.

To reduce the potential clinical risks, significant research has been conducted to minimize both the aforementioned factors. However, the quality of X-ray images still depends significantly on cardiac motion artifacts and various types of noise, complicating PCI image guidance and the accurate diagnosis of coronary arteries by operators based on X-ray images.

Moreover, some parts of coronary arteries may be completely blocked, particularly in chronic total occlusion (CTO) cases, which further complicate PCI and diagnosis as they are not visible in X-ray images. Therefore, PCI procedure tends to be highly dependent on the intuition and experience of the operator.

Image segmentation processing is helpful for image-based diagnosis and procedure planning. However, segmentation models may fail to segment regions where blocked parts are not visible in X-ray images. For instance, in the X-ray image of a normal person depicted in Fig. 1(a), the predicted segment of the coronary artery does not include any broken regions (Fig. 1(b)). In contrast, X-ray images of patients with CAD or CTO (Fig. 1(c)) predicted segments with broken regions (Fig. 1(d)). The reconstruction of these broken regions is an important problem to make blocked regions completely visible, thereby facilitating PCI.

To this end, several registration-based approaches for procedure guidance have been proposed to reduce the uncertainty inherent in 2D X-ray images via feature matching between 2D X-ray images and three-dimensional (3D) CT images [4], [5], [6], [7], [8], [9], [10]. In [4], statistical motion models of coronary arteries based on 4D CT angiography (CTA) were introduced, and a 2D/3D+t coronary artery registration method using motion models based on cardiac

and respiratory information was proposed in [5]. Additionally, Pernus et al. [6] used intensity gradients of 2D X-ray images to match 3D vascular geometry and Zhu et al. [7] introduced a matching method based on iterative closest graphs using coarse-to-refine vessel matching for rigid and non-rigid transformation. Recently, convolutional neural network (CNN) models have been used to extract the central lines of coronary arteries in 2D X-ray images and energy function-based 3D deformation has been utilized for real-time registration [8], [9], [10].

However, these methods are applicable only if an additional 3D CT image is scanned from each patient. Further, registration-based methods are significantly dependent on segmentation performance during the extraction of the central lines of coronary arteries from both 2D X-ray and 3D CT images. The segmentation tasks are known to be challenging even by themselves. Moreover, even after extracting central lines successfully from both modalities, registration remains challenging and computationally expensive, making these methods difficult to apply in emergency situations.

We assume that the ideal approach minimizes the dependency on external information, such as CT images. In our previous study [11], we proposed a re-connection method based on detection and connection of local vascular geometries without any external information. However, the performance of the proposed method depended on the performance of the initial segmentation model. Further, the local broken regions were sometimes over-compensated for, even when the re-connection task was performed successfully.

In this study, instead of reconnecting broken segments in the absence of vascular information, we design a novel generative model to reconstruct broken regions robustly without any external dependence. The proposed method is based on generative adversarial networks (GANs) [12]. A GAN consists of two components—a generator G and a discriminator D. They are trained in an adversarial manner to synthesize realistic image values and have achieved great success in diverse tasks, such as image synthesis [13], image translation [14], image inpainting [15], [16], [17], and even medical imaging reconstruction [18], [19], [20].

Among these tasks, we focus on image inpainting, which involves the restoration of impaired images or those containing certain naturally erased objects. We assume that the broken parts of coronary arteries in X-ray images can be naturally reconstructed via inpainting. The generative model realistically restores broken parts without using any external information if it is trained sufficiently on normal vascular structures.

We propose a fully automatic method based on GAN to restore broken regions of coronary arteries. The contributions of this paper can be summarized as follows:

- To the best of our knowledge, the proposed method is the first approach that uses a GAN-based model to reconstruct blocked regions directly without depending on any external modality.

- For the optimal performance, we first introduce the vesselness-based objective function (vesselness-loss) specifically tailored for elongated objects.
- Novel multi-scale aggregation block (MAB) networks are proposed for multi-scale image encoding that considers both global and local context information.
- The proposed method outperforms the state-of-the-art methods in terms of three quantitative evaluation metrics—peak signal-to-noise ratio (PSNR), mean squared error (MSE), and structural similarity index measure (SSIM), and also generates the vessel-likes images in qualitative terms.
- The proposed method can be utilized for image-based procedure guidance as it yields quantitatively and qualitatively plausible vessel paths even in the cases of severe CAD patients.

The remainder of this paper is organized as follows. Related works are outlined in Section II. The fully automatic vessel reconstruction process is introduced in Section III. Next, the extensive set of experiments performed to evaluate the proposed method and their results are described in Section IV. Finally, related discussions and conclusions are presented in Sections V and VI, respectively.

II. RELATED WORKS
A. CLASSICAL INPAINTING METHODS
Traditional inpainting methods fill holes in images based on existing content. These methods can be categorized into two main classes—diffusion-based [22], [23] and patch-based methods [24], [25], [26]. Diffusion-based methods fill missing regions using undamaged neighboring information, whereas patch-based methods search for patches similar to missing holes in the background or external databases and use them to complete the image.

Although simple cases involving stationary backgrounds or repeating patterns, for example, can be inpainted quite effectively using such traditional methods, they fail in complex cases owing to lack of high-level structural understanding of the entire images. Thus, the inpainted results are visually inconsistent with their surroundings, implausible, and unsatisfactory.

B. LEARNING-BASED INPAINTING
Natural Image Inpainting In recent years, deep-learning-based solutions [15], [16], [27], [28], [29], [30] have exhibited superior performance compared to classical non-learning-based approaches, and have been demonstrated to be capable of synthesizing more visually plausible and satisfactory content even in complex scenes. By exploiting large-scale data in either supervised or unsupervised adversarial manners [12], rich texture patterns are learned by such methods, and missing or damaged regions are filled with reasonable content based on features located at great distances from the target regions.

The context encoder [15] was the first learning-based approach introduced for this task—it utilizes an encoder-decoder architecture with adversarial loss and Euclidean distance-based loss. Iizuka et al. [27] further improved synthesized image quality by incorporating dilated convolution into their encoder-decoder architecture and utilizing local and global discriminators, which ensure semantic coherence between original regions and generated regions. Subsequently, the contextual attention module [28] was proposed to capture long-range contexts, further improving performance.

Meanwhile, coarse-to-refine frameworks were devised in [28], [29], and [31] where each stage utilizes a novel auto-encoder network. A model similar to that developed in [31] was introduced in [21]—each employs a two-stage training scheme, but the latter uses an attention computing module (ACM) and attention transfer module (ATM) to generate attention scores and create aggregated residuals to synthesize high-quality images.

In conjunction with two-stage pipelines, additional edge information has been exploited to generate more realistic images [16]. A two-stream network has also been proposed by guiding and constraining texture and structure in a balanced manner to achieve more detailed image synthesis [32]. In both of the aforementioned methods, the edges of holes are generated via specific generators, instead of via propagation to another objective function.

Medical Imaging Inpainting Early research on medical imaging inpainting [33], [34], [35] focused on removing artifacts or anomalies for targeting tasks, e.g., image segmentation and registration. Deep learning has shifted the paradigm of medical imaging inpainting.

In two successive studies, Armanious et al. proposed adversarial learning frameworks [18], [36] for MR images that synthesize missing content in rectangular and arbitrarily shaped regions. Meanwhile, [20] used information on edge, structure, multi-scale residual blocks, and multi-scale loss to synthesize more realistic CT and MRI images.

C. MULTI-SCALE FUSION
Multi-scale fusion blocks have been used for single-image super-resolution [37], [38]. Also, they have been utilized in conjunction with self-guided regression loss to generate visually coherent images [39]. Subsequently, [17] proposed Aggregated contextual Transformation (AOT) blocks that aggregate contextual transformations from various receptive fields.

D. MULTI-SCALE VESSEL EXTRACTION
Frangi et al. [40] devised a method to distinguish continuous edges and tubular structures, such as vessels, using Hessian matrices. Additionally, [40] and [41] proposed multi-scale Frangi filters, in which the eigenvalues of Hessian matrices are emphasized for vessel-like structures. This filter has been widely used for vessel segmentation [42], [43]. In particular, [42] demonstrated the effectiveness of the

TABLE 1. The methodologies, pros and cons of baseline methods comparing our proposed methods.

Methods	Methodologies	Pros	Cons
CE [15]	Proposes Context Encoders that utilizes context to restore the missing region.	Demonstrates its superiority by adopting to pre-training.	Still blurry and unsatisfactory in some cases.
IPA [18]	Optimizing using CasNet generator, global and patch discriminator and VGGNet-based feature extractor in an adversarial fashion.	Being able to handle arbitrary missing regions and produces consistent textures and structures.	Requires a high computational cost from the generator.
HR [21]	High-resolution inpainting with lightweight gated convolution and Contextual Residual Aggregation.	Being able to fill the large holes using a generator with comparatively less parameters.	Difficult to optimize 2-stage framework and computationally expensive to compute multiple attention transfer.
AOT [17]	Uses AOT (Aggregated Contextual Transformation) block and Soft Mask-Guided PatchGAN discriminator, which uses soft labels generated by Gaussian filtering.	Easier to optimize and computationally more efficient than 2-stage approaches.	Tends to encode predefined gridding patterns rather than patterns for context reasoning.

Notes: The outlines of baseline methods comparing our proposed methods.

Frangi-based filter in multi-scale vessel extraction using fundus images.

As is evident from the preceding discussion, the reconstruction of broken regions in X-ray images using Hessian-guided objective functions (vesselness-loss) to guide generative models implicitly has not been researched yet.

Table 1 describes the methodologies, advantages, and disadvantages of prior similar works used in our experiments.

III. METHODS

Ideally, in X-ray images of normal cases, coronary arteries do not contain any broken regions. However, in X-ray images of patients with coronary artery disease (CAD), coronary arteries may appear to be broken into several segments.

Let the coronary regions and segments be denoted by C and $c_i \in C$, respectively, and the broken parts and the other intact parts be denoted by C^{broken} and C^{vessel}, respectively. Then, the ideal coronary artery can be denoted by $C = \{C^{\text{vessel}}\}$, whereas the broken coronary artery can be denoted by $C = \{C^{\text{broken}} \cup C^{\text{vessel}}\}$.

It is worth noting that coronary arteries are very thin and, thus, sparsely represented in raw data. If random stroke masking is used on data following previous studies [16], [32] the inpainting model may suffer from being significantly affected by background information, resulting in sub-optimal performance.

Meanwhile, in our previous work [11], we demonstrated that the model proposed therein was capable of defining a local broken region, c_i^{broken}, given the region, C^{vessel}. In other words, local broken regions, c_i^{broken}, can be extracted automatically and directly. Thus, for the first stage of our framework, we leverage our previous method to detect local broken regions and avoid the aforementioned issue.

Using this technique, we extract a square-shaped patch X_i^{patch} centered on c_i^{broken}. We take the size of X_i^{patch} to be sufficiently large to utilize probable vessel direction flow information (local) and neighboring area information (global). Also, we allow the vessel information in the ROI

X_i^{patch} to be completely blocked, rendering the broken parts completely invisible. This situation is described by Eq. 1:

$$X_i^{\text{blocked}} = X_i^{\text{patch}} \odot (1 - \text{M}) + \text{M} \qquad (1)$$

where M has the same shape as X_i^{patch} and consists of values, 1 (corresponding to the region of interest) and 0 (corresponding to the external original region), where $\odot$ represents the Hadamard product.

After G reconstructs X_i^{blocked}, it is blended with the original input of X_i^{patch} using the mask, M, to preserve the external original region that ought to be consistent during the reconstruction process. This is described by Eq. 2.

$$\tilde{X}_i^{\text{blocked}} = X_i^{\text{patch}} \odot (1 - \text{M}) + G(X_i^{\text{blocked}}) \odot \text{M} \qquad (2)$$

The final objective of this study is to reconstruct the broken parts realistically such that $\tilde{X}_i^{\text{blocked}}$ is indistinguishable from X_i^{patch}.

In the following parts of this section, we present the novel method and compare its architecture and objective function with those of existing state-of-the-art methods. First, we describe the novel MAB generator and discriminator. Next, various objective functions and the novel vesselness-loss objective function used in our experiments are described in detail. Fig. 2 provides an intuitive illustration of the training and testing pipelines, and Fig. 3 depicts a core module of our MAB network.

A. ARCHITECTURES

1) GENERATOR

To reconstruct broken regions containing potentially important vessel information, information from both adjacent and distant contexts should be considered. The proposed model synthesizes the broken region using both local and global information. An overview of the generator is illustrated in Fig. 2. The generator is a single-stage network comprising an encoder, a stack of novel multi-scale aggregation blocks (MAB), and a decoder. It receives an image and a mask indicating the missing pixels values, and outputs a restored image. The details are described below:
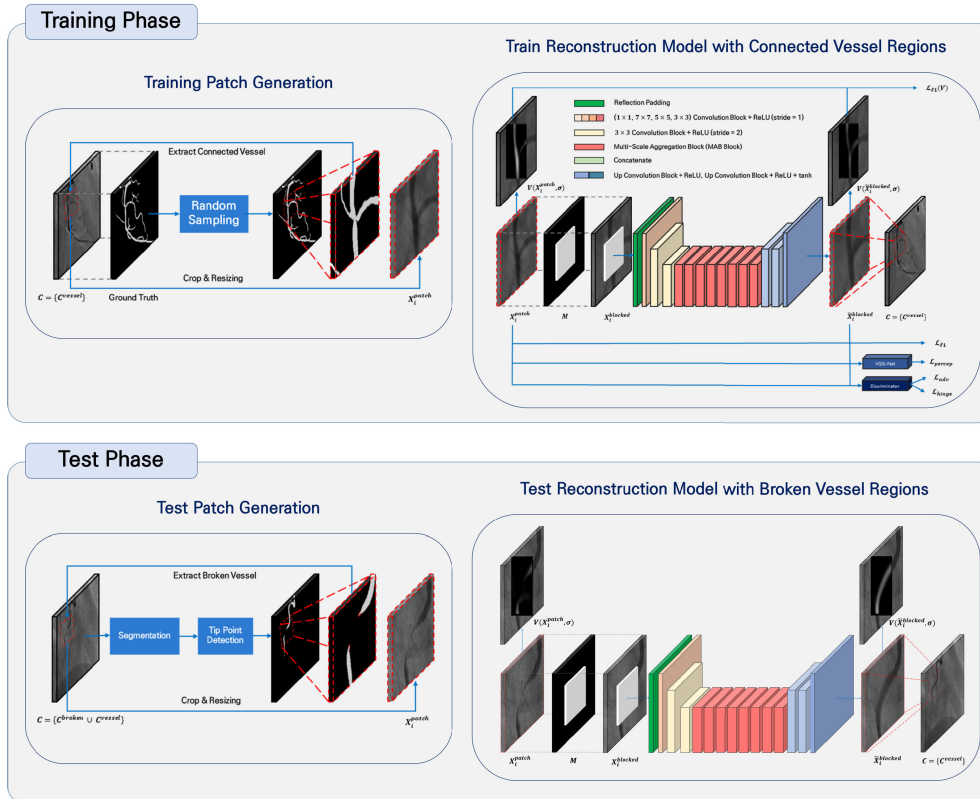
FIGURE 2. Fully automatic vessel information reconstruction workflow. **Training Phase** X-ray patches X_i^{patch} containing ideal connected coronary arteries $C = \{C^{vessel}\}$ are extracted via random sampling of their corresponding ground truths. Next, extracted patches are synthetically blocked or damaged to simulate various situations that may be encountered in actual clinical situations. Then, blocked or damaged patches are transmitted into models to enable them to learn to reconstruct broken vessels into ideal ones. In particular, in the reconstruction process, various loss functions are used; to force the model to focus on sparse vessel regions explicitly, the novel vesselness-loss $\mathcal{L}_{\ell 1(V)}$ is applied, minimizing differences between vesselness probability maps $V(X_i^{patch}, \sigma)$ and $V(\tilde{x}_i^{blocked}, \sigma)$, obtained from the region of interest (ROI) (X_i^{patch}) and the model output ($\tilde{x}_i^{blocked}$) at scale σ, respectively. **Test Phase** In this phase, extreme cases, such as those involving coronary artery disease, that can lead to the absence of vessel information, are simulated. Our previous methods [11] perform segmentation masking on the cases, followed by tip point detection to find X_i^{broken} containing broken coronary arteries $C = \{C^{broken} \cup C^{vessel}\}$ with a sufficient size to cover broken coronary areas C_i^{broken}. Then, our trained model receives those areas and outputs results $\tilde{x}_i^{blocked}$ that are similar to the ideal ones C^{vessel}. This phase is rephrased as follows. Corresponding to the broken coronary artery ($C = \{C^{broken} \cup C^{vessel}\}$), the model aims to reconstruct it from C^{broken} to C^{vessel} so that $C^{broken} = \phi$. In other words, the broken coronary artery ($C = \{C^{broken} \cup C^{vessel}\}$) is reconstructed to be similar to the ideal coronary artery ($C = \{C^{vessel}\}$).

Encoder & Decoder We use an encoder-decoder structure based on vanilla convolutions for a single-refinement inpainting network, where both the encoder and decoder are composed of three convolutional layers. Following previous studies [28], [30], we do not use any kind of normalization layer to avoid color shift.

MAB Features compressed and propagated by the encoder are transmitted to a stack of MAB blocks. Inspired by previous studies that used large kernels [44], [45], [46] and multi-scale fusion blocks [17], [32] for more effective representation learning and image restoration, respectively, we construct a simple but novel MAB block, bridging two branches of previous studies. A diagram of the block is presented in Fig. 3.

To design the basic structure of the block, we adopt a ***split-transform-merge*** strategy. ***Split***: The input from the previous layer or encoder is split and propagated to multiple blocks with varying kernel sizes of 1, 3, 5, and 7. ***Transform***: As mentioned, both adjacent and distant contexts are important in inpainting. For adjacent contexts, features from convolutional blocks of kernel sizes 3 and 5 are concatenated. For distant contexts, features from blocks of kernel sizes 1 and 7 are concatenated. The two concatenated features are transmitted to a convolutional layer with a kernel size of 1 to enhance expressiveness and efficiency. The first branch is effective at capturing and recovering global background information as a convolutional layer with large kernel size (7×7) is integrated with a pixel-wise convolutional layer (1×1), whereas the second branch handles vessel-specific local information. ***Merge***: The blocks are concatenated again for integration and transmitted to a convolutional layer with a kernel size of 3. Additionally, inspired by the great success
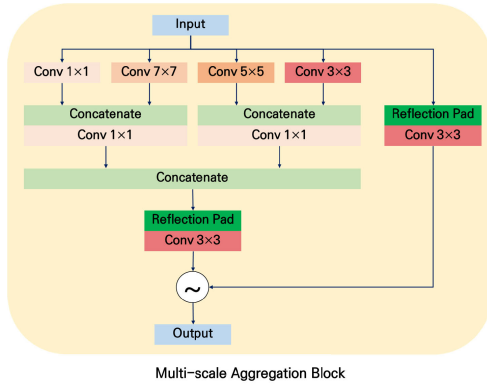
FIGURE 3. Multi-scale aggregation block (MAB) incorporates convolutional layer with large kernel size into multi-scale fusion blocks.

of ResNet [47], we add information flow from another convolutional layer with kernel size 3 to facilitate training and amplify ensemble effect [48] during restoration.

Here, we adopt the gated residual connection strategy, proposed in [17], weighted sums are obtained using gated values in place of simple residual summation.

2) DISCRIMINATOR

The discriminator used in this task is obtained from Spectral-Normalized Markovian PatchGAN (SN-PatchGAN) [49], [50], following previous studies [17], [21], [30]. It is simple in formulation but enables fast and stable training, thereby producing high-quality samples.

B. LOSS FUNCTIONS

To synthesize images realistically while preserving structure, the generator, G, and the discriminator, D, should be trained simultaneously in an adversarial manner. In this section, we introduce various loss functions used for this purpose in detail.

1) HINGE LOSS

Lim and Ye [51] proposed a loss function based on a soft-margin support vector machine (SVM) linear classifier, which utilized a hyperplane that maximizes the margin between two distributions. It has been verified that this method produces more stable training and constraints the occurrence of mode collapses. Eq. (3) and (4) describe the generator and discriminator, respectively:

$$\mathcal{L}_G = -\mathbb{E}_{x \sim P_{\tilde{X}^{\text{blocked}}}}[D(\tilde{X}_i^{\text{blocked}})] \tag{3}$$

$$\mathcal{L}_D = E_{x \sim P_{X^{\text{patch}}}}[ReLU(1 - D(X_i^{\text{patch}})] \\ + E_{x \sim P_{\tilde{X}^{\text{blocked}}}}[ReLU(1 + D(\tilde{X}_i^{\text{blocked}})] \tag{4}$$

2) INDUCED RECONSTRUCTION OF ENTIRE CONTEXT INFORMATION

A contextual loss function is required to generate $\tilde{X}_i^{\text{blocked}}$ that is semantically close to the actual image X_i^{patch}. For this purpose, we use the $\ell1$ loss function instead of the $\ell2$

loss function as the former converges more easily, preserves sharp details more robustly and, the most importantly, is less vulnerable to yielding blurry results. Using $\tilde{X}_i^{\text{blocked}}$ and X_i^{patch}, we define Eq. (5) as follows:

$$\mathcal{L}_{\ell 1} = ||\tilde{X}_i^{\text{blocked}} - X_i^{\text{patch}}||_1 \tag{5}$$

3) STRUCTURAL SIMILARITY INDEX LOSS

The structural similarity index measure (SSIM) [52], as shown in Eq. (6), is a metric that is used to measure image quality. Unlike $\mathcal{L}_{\ell 1}$ in Eq. (5), which simply compares pixel-wise differences, we optimize our models in terms of SSIM, which uses both luminance and contrast. Thus, SSIM represents human visual perception more faithfully. This index is determined based on correlation coefficients—with high values corresponding to generated images that are more qualitatively plausible to the naked human eye.

$$SSIM(X, Y) = \frac{2\mu_X \mu_Y + C_1}{\mu_X{}^2 + \mu_Y{}^2 + C_1} \cdot \frac{2\sigma_{XY} + C_2}{\sigma_X{}^2 + \sigma_Y{}^2 + C_2} \tag{6}$$

For simplicity, we denote $\tilde{X}_i^{\text{blocked}}$ and X_i^{patch} by X, Y respectively. Moreover, μ and σ denote average and variance, respectively, $C_1 = (k_1 L)^2$ and $C_2 = (k_2 L)^2$ are used, the pixel intensity is denoted by L, and the constants, k_1, k_2, are set to 0.01 and 0.03, respectively. Finally, the SSIM loss can be defined using Eq. (7).

$$\mathcal{L}_{SSIM} = 1 - SSIM(X, Y) \tag{7}$$

4) PERCEPTUAL LOSS

Perceptual loss [53] has been widely used in low-level vision tasks, e.g., inverse problems such as image synthesis and image reconstruction, owing to its ability to generate more visually plausible results. It encourages similarities between the target and synthesized images in a feature space of pre-trained VGGNet [54] models. For this purpose, we adopt the relu5_1 layer of VGG19.

$$\mathcal{L}_{per} = \frac{1}{C_i H_i W_i} ||\phi_i^{vgg}(\tilde{X}_i^{\text{blocked}}) - \phi_i^{vgg}(X_i^{\text{patch}})||_1 \tag{8}$$

where $\phi_i^{vgg}(\cdot)$ denotes the feature map at the i-th layer of VGG19, and C_i, H_i and W_i denote elements of that feature map, indicating channel, height, and width, respectively. We also use $\ell1$ loss for the reasons described earlier.

5) ENFORCED RECONSTRUCTION OF VESSEL INFORMATION

In Eq. (5), to generate a semantically close $\tilde{X}_i^{\text{blocked}}$ while preserving the contextual information of X_i^{patch}, the intensity difference at every pixel between $\tilde{X}_i^{\text{blocked}}$ and X_i^{patch} was calculated. However, calculating the intensity difference at all pixels in the image is so vague because the model should reconstruct the broken vessels focusing on the characteristics of thin and sparsely existing blood vessels. This difficulty necessitates the development of a new objective function that

preserves contextual information by concentrating only on the vascular region, occupying a very small area in a thin and elongated form in the image.

In this study, in conjunction with Eq. (5), which is used to preserve the contextual information of the entire image, a Hessian-based loss is used to preserve the vascular contextual information by focusing on the local vascular region in the image.

The Hessian matrix, **H**, represents the curvature characteristics of a function by Eq. (9), and its elements are second-order derivatives and a linear transformation that makes a certain bowl-shaped function geometrically more convex or concave. Based on the eigenvalues, λ, and eigenvectors, v, of the Hessian matrix, the degree of change in the linear transformation can be ascertained. The eigenvector represents a direction vector with a large function curvature, whereas the eigenvalue represents the magnitude of the curvature of the function in the direction of the corresponding eigenvector.

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 I(x,x)}{\partial x^2} & \frac{\partial^2 I(x,y)}{\partial xy} \\ \frac{\partial^2 I(y,x)}{\partial yx} & \frac{\partial^2 I(y,y)}{\partial y^2} \end{bmatrix} \quad (9)$$

Based on these characteristics of the Hessian matrix, several approaches have been proposed using the difference between its eigenvalues as a threshold for the complete segmentation of vascular regions from the background region [40], [41], [42], [43]. Among them, Frangi et al. [40] developed the most representative Hessian-based vascular segmentation technique based on the concept of vesselness probability. Vesselness probability represents the probability that a specific pixel belongs to a blood vessel and can be derived using the ratio, R, between the magnitudes of the two eigenvectors and the corresponding eigenvalues, λ_1 and λ_2, obtained from the Hessian matrix. Then, using the ratio $R_b^2 = \frac{|\lambda_1(p,\sigma)|}{|\lambda_2(p,\sigma)|}$ and $S = \sqrt{\lambda_1^2 + \lambda_2^2}$, Hessian-matrix-based vesselness probability at a pixel $p = (x, y)$ can be obtained in multiple scales using Equation (10).

$$v(p, \sigma) = \begin{cases} 0 & if \lambda_2 > 0, \\ \exp(-\frac{R_b^2(p,\sigma)}{2\beta^2})(1 - \exp(-\frac{S^2(p,\sigma)}{2C^2})) \end{cases} \quad (10)$$

In equation (10), R and S indicate the blobness and structuredness, respectively, when $|\lambda_1(p, \sigma)| < |\lambda_2(p, \sigma)|$ at x with scale s. Then, if we denote the set of pixels in the image by X, the vesselness probability map of the entire image I on scale σ can be defined to be $V(I, \sigma)$.

However, the obtained vesselness probability based on the Hessian matrix, **H**, varies with respect to the scale, σ, of the matrix, as illustrated in Fig. 4. When σ is small, micro-vessels and false positives (FP) are segmented because the probability is high in the vessel-like structure. On the other hand, when σ is large, micro-vessels cannot be segmented satisfactorily as the probability is high only in the clear vessel-like structure,

thereby reducing the number of FPs. Thus, a trade-off exists between σ and the segmentation accuracy of vessels.

In this study, FPs generated during the reconstruction of blood vessels, including micro-vessels, are minimized by taking $\sigma = 3-7$ to cover both micro-vessels and thick blood vessels elaborately. Then, using $\tilde{X}^{patch}$, the generator, G, reconstructs thin and sparse blood vessels in X^{patch} by preserving the information of the vessels by optimizing $\ell 1$ based on Equation (11).

Then, using $\tilde{X}^{patch}$, generator G can reconstruct thin and sparsely existing blood vessels in X^{patch} by preserving the information of the vessels through optimizing $\ell 1$ based vesselness-loss as Eq. (11).

$$\mathcal{L}_{\ell 1(V)} = ||V(\tilde{X}_i^{blocked}, \sigma) - V(X_i^{patch}, \sigma)||_1 \quad (11)$$

C. LINEAR COMBINATION OF LOSS FUNCTIONS

In the previous sections, five loss functions are introduced to improve the quality of the synthesized vessel information. They are now summed linearly with weighted constants. Eq. (12) and (13), describe the objective terms for the discriminator, D, and generator, G, respectively.

$$\mathcal{L}_{D_{total}} = \lambda_1 \mathcal{L}_{hinge} \quad (12)$$

$$\mathcal{L}_{G_{total}} = \lambda_1 \mathcal{L}_{hinge} + \lambda_2 \mathcal{L}_{\ell 1} + \lambda_3 \mathcal{L}_{SSIM} + \lambda_4 \mathcal{L}_{per} + \lambda_5 \mathcal{L}_{\ell 1(V)} \quad (13)$$

IV. EXPERIMENTS AND RESULTS

In this section, we describe extensive experiments to quantitatively and qualitatively evaluate the performance of the proposed method based on reconstructed broken parts. First, the details of the dataset used for the experiments are outlined in Section IV-A. For fair comparison, we use popular evaluation metrics, which are introduced in Section IV-B. Moreover, the reconstruction performance is evaluated both quantitatively and qualitatively using synthesized images to simulate various cases encountered in real-world clinical sites; this evaluation is presented in Section IV-C. The reconstruction performance is then evaluated under the assumption of complete blockage of the vessel, as discussed in Section IV-D. Additionally, the performance of the proposed fully automated process is demonstrated by presenting the result of each step in Section IV-E. Finally, ablation studies performed to demonstrate the effectiveness of our novel vesselness-loss are described in Section IV-F.

All of the following experiments are conducted in an Ubuntu 18.04 environment with 64 AMD EPYC 7513 32-Core CPU Processors, 1 TB RAM, and NVIDIA®RTX A6000 GPU. We use Pytorch version 1.10 as the main deep learning framework.

A. DATASET

The proposed methods were trained and evaluated using a dataset, enrolled clinically stable adult patients from September 2015 to February 2016 at Severance Cardiovascular Hospital who underwent clinically indicated ICA.
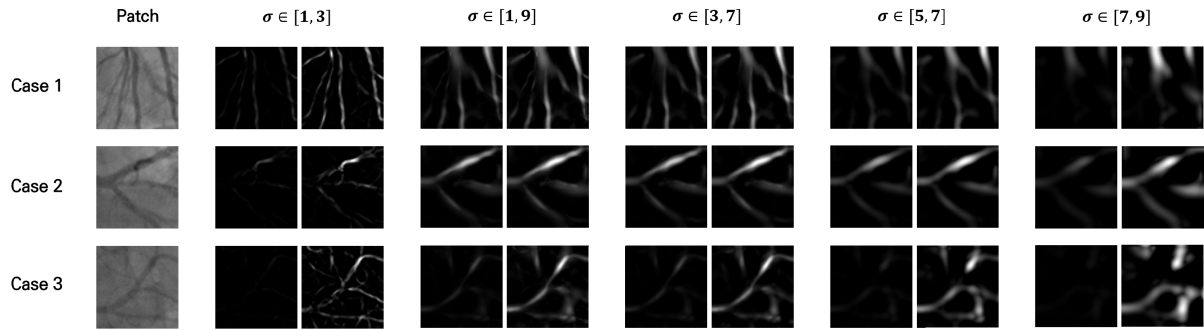
FIGURE 4. Results of Hessian-based vesselness probability map and normalization depending on the scale, σ, indicating a trade-off relationship. When σ is low, vessel-like structure region has high vesselness probabilities, and thus thin vessels are well segmented; however, vessel-like false positives are segmented together. In contrast, when σ is high, the vesselness probabilities are high only in thick vessel-like structures, and thus thin vessels are not well segmented, and false positives are reduced.

Institutional Review Board (Severance Hospital, IRB Number 1-2017-0031) approval was obtained for this retrospective study and informed consent was waived. The dataset comprises 3,136 2D X-ray angiography images captured from 85 patients, and corresponding ground truth (GT) coronary artery regions marked by clinical experts with more than 5 years of experience. No significant difference exists in intensity or structure among the images in the dataset. Each image is reconstructed as a 512 × 512 image with a size between 0.28 mm × 0.28 mm and 0.36 mm × 0.36 mm for each pixel.

The models are trained to restore blood vessels more robustly by randomly sampling pixels corresponding to the coronary region in the GT to generate 128 × 128 patches centered on the pixel. When a 64 × 64-sized blocked mask, sufficient to cover the coronary artery, is positioned at the center of a patch where actual blood vessels are restored, only those 128 × 128 patches depicting more than 30% of the coronary artery in the same region in the GT are used for learning. The dataset is divided into training and testing sets—1,994 images from 50 patients are used for training and 1,142 images from 35 patients are used for evaluation. Then, via patch generation, 20,752 patches extracted from the 1,994 images are used for training, and 11,808 patches extracted from the 1,142 images are used for evaluation. The details of the dataset used in the experiment are listed in Table 2.

B. EVALUATION METRICS
For the quantitative evaluation of the proposed method with other state-of-the-art counterparts, three major metrics are used.
- **PSNR**: PSNR is a frequently used metric to evaluate the amount of quality loss in a generated image. The lower the degree of quality loss, the higher the PSNR value. PSNR is defined as the log of the value obtained by dividing the square of the maximum value, *MAX*, of a pixel (peak signal) that can be expressed in the image by the mean squared error (MSE), MSE, as given by Equation (14).

$$PSNR = 10 \cdot \log_{10} \frac{MAX^2}{MSE} \qquad (14)$$

TABLE 2. The Details of dataset used for experiments.

Size	Dataset	#Patient	#Frame	# Frame/Patient
512 × 512	Total	85	3,136	[19, 66]
	Training	55	1,994	[19, 66]
	Test	30	1,142	[22, 61]
128 × 128	Total	85	32,560	[1, 39]
	Training	55	20,752	[1, 38]
	Test	30	11,808	[1, 39]

Notes: The dataset was provided by Severance Hospital, Yonsei University College of Medicine, South Korea, comprising 3,136 2D X-ray angiography images from total of 85 patients and the corresponding GT.

- **MSE**: It is obtained by dividing the square of the error of each sample by the number of samples. During image quality evaluation, this value indicates the mean square of the intensity difference for each pixel—the lower the loss, the lower the MSE value. If y_i denotes the actual value, and $\hat{y}_i$ denotes the predicted value, MSE is given by Equation (15).

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad (15)$$

- **SSIM**: SSIM can be used not only as an additional loss function to generate an image, as described in Section III-B3, but also as one of the main evaluation metrics used to measure luminance, contrast, and structural similarity between original and generated images, as given by Equation (6).

C. EVALUATION IN THE CASE OF DAMAGED VASCULAR INFORMATION
The proposed method is first evaluated on cases involving damaged vessel information pertaining to non-severe diseases, e.g., with weak stenosis or inherent limitations of X-ray images. To simulate various cases encountered by real-world clinicians, arbitrary damaged images are synthesized by adding Gaussian blur or noise to the vessel information in intact X-ray images.

We denote the Gaussian blur kernel by B and the additional Gaussian noise added to each pixel by N. Then, we apply

TABLE 3. Quantitative comparison of various methods on the reconstruction of synthesized arbitrary damaged images.

Degree of arbitrary damaging	Weakly damaged			Moderately damaged			Strongly damaged		
Metrics	PSNR ↑	MSE ↓	SSIM ↑	PSNR ↑	MSE ↓	SSIM ↑	PSNR ↑	MSE ↓	SSIM ↑
CE [15]	33.104	0.231	0.906	32.779	0.247	0.891	31.223	0.342	0.879
CE + $\mathcal{L}_{SSIM}$	33.263	0.208	0.899	32.645	0.252	0.889	31.150	0.344	0.876
CE + $\mathcal{L}_{\ell1(V)}$	33.609	0.195	0.902	**32.785**	**0.246**	**0.891**	**31.649**	**0.317**	**0.881**
CE + $\mathcal{L}_{\ell1(V)} + L_{SSIM}$	**34.051**	**0.178**	**0.907**	32.504	0.257	0.888	31.405	0.330	0.879
IPA [18]	34.072	0.181	0.902	33.034	0.233	0.894	32.870	0.270	0.899
IPA + $\mathcal{L}_{SSIM}$	35.277	0.133	0.922	33.376	0.210	0.906	32.840	0.254	0.900
IPA + $\mathcal{L}_{\ell1(V)}$	35.291	0.137	0.922	**33.847**	**0.199**	**0.908**	**33.122**	**0.238**	**0.901**
IPA + $\mathcal{L}_{\ell1(V)} + L_{SSIM}$	**35.477**	**0.126**	**0.925**	33.711	0.200	0.907	32.895	0.249	0.899
HR [21]	34.134	0.179	0.903	32.692	0.253	0.887	31.983	0.302	0.881
HR + $\mathcal{L}_{SSIM}$	34.173	0.176	0.901	32.788	0.249	0.889	32.078	0.297	0.882
HR + $\mathcal{L}_{\ell1(V)}$	**34.338**	**0.168**	**0.904**	32.786	0.248	0.887	**32.508**	0.299	0.881
HR + $\mathcal{L}_{\ell1(V)} + L_{SSIM}$	34.274	0.173	0.903	**32.940**	**0.241**	**0.891**	32.108	**0.294**	**0.883**
AOT [17]	35.418	0.130	0.917	33.833	0.197	0.901	32.763	0.251	0.894
AOT + $\mathcal{L}_{SSIM}$	35.332	0.131	0.919	33.744	0.200	0.905	32.773	0.253	0.896
AOT + $\mathcal{L}_{\ell1(V)}$	**35.553**	**0.126**	0.918	33.804	0.194	0.902	32.813	0.253	0.892
AOT + $\mathcal{L}_{\ell1(V)} + L_{SSIM}$	35.503	0.129	**0.922**	**33.899**	**0.193**	**0.908**	**32.957**	**0.245**	**0.896**
MAB	35.677	0.121	0.919	33.852	0.193	0.902	33.206	0.227	0.898
MAB + $\mathcal{L}_{SSIM}$	35.653	0.121	0.919	33.827	0.193	0.902	33.016	0.234	0.896
MAB + $\mathcal{L}_{\ell1(V)}$	35.669	0.120	0.918	33.990	0.186	0.904	33.127	0.231	0.897
MAB + $\mathcal{L}_{\ell1(V)} + L_{SSIM}$	**35.741**	**0.119**	**0.921**	**34.014**	**0.186**	**0.905**	**33.246**	**0.225**	**0.899**

Notes: As evidenced in the table, our MAB networks achieve highest performance over other strong baseline networks (CE [15], IPA [18], HR [21], and AOT [17]). Furthermore, it is evident that utilizing the novel vesselness-loss contributes to performance improvement when it is used with other loss functions, e.g., SSIM, and results could often be synergistic.

Gaussian blur or Gaussian noise n times, which can be denoted by B^n and N^n, respectively. If $\otimes$ represents the dot product, the damaged vascular regions, X_i^{damaged}, in arbitrary levels depending on B^n and N^n can be defined by Eq. 16.

$$X_i^{\text{damaged}} = (X_i^{\text{patch}} \odot (1 - M) + ((X_i^{\text{patch}} \odot M) \otimes B^n + N^n) \tag{16}$$

Then, the remaining non-damaged regions, which may be affected during the noise-reduction process, are preserved by defining the output, $\tilde{X}_i^{\text{damaged}}$, of the generator, G, using Eq. 17.

$$\tilde{X}_i^{\text{damaged}} = X_i^{\text{patch}} \odot (1 - M) + G(X_i^{\text{damaged}}) \odot M \tag{17}$$

For fair comparison between the proposed MAB network and other state-of-the-art networks [15], [17], [18], [21], the loss functions is taken to be identical to that of the baseline network described in the original papers. Experimental results reveal that the proposed MAB network outperforms the other methods in three experimental environments—weakly damaged (B^1, N^1), moderately damaged (B^3, N^1), and strongly damaged (B^5, N^1). Moreover, the following constant values of the loss terms yield the best results: $\lambda_1 = 1$, $\lambda_2 = 256$, $\lambda_3 = 1$, $\lambda_4 = 128$ and $\lambda_5 = 256$.

As shown by the quantitative comparison outlined in Table 3, the proposed MAB network exhibits the highest PSNR, MSE, and SSIM scores of 35.666, 0.121, and 0.919, respectively, over strong baseline networks (CE, IPA, HR, AOT, and MAB) in weakly damaged settings. The proposed MAB network also exhibits the best PSNR, MSE, and SSIM scores of 33.852, 0.193, and 0.902, respectively, in moderately damaged settings. Even in the strongly damaged case, the MAB network exhibits the best performance in

terms of PSNR and MSE, with scores of 33.206 and 0.227, respectively. However, its SSIM score of 0.898 is slightly lower than that of 0.899 obtained by IPA; this is because SSIM can be higher in blurry images [55].

Next, we evaluate the performance improvement of the vessel-specific restoration process achieved using our proposed vesselness-loss function, $\mathcal{L}_{\ell1(V)}$. Additional experiments are conducted with the regularization of the baseline networks and the proposed MAB network using $\mathcal{L}_{\ell1(V)}$ or L_{SSIM} or both. As shown in Table 3, adding L_{SSIM} does not always lead to improved vessel-specific reconstruction performance. On the contrary, optimization using $\mathcal{L}_{\ell1(V)}$ contributes to improvement in vessel reconstruction—every baseline model exhibits its highest quantitative results in all arbitrarily damaged settings. Further, training networks using $\mathcal{L}_{\ell1(V)}$ in conjunction with L_{SSIM} results in higher quantitative results in some cases, implying that improved results can be expected when the model is optimized along with other objective functions.

The reconstructed results of each baseline network when trained with $\mathcal{L}_{\ell1(V)}$ or both $\mathcal{L}_{\ell1(V)}$ and L_{SSIM} are presented in Fig. 5. The proposed MAB network achieves the highest performance in vessel reconstruction, as demonstrated by the data in Fig. 5 and Table 3.

The proposed network is capable of reconstructing the most robust results compared to the baseline networks that are consistent with neighboring pixels and the most similar to the intact X-ray images visually. In comparison, the other baselines [15], [18], [21] generate blurry results with the broken parts in several regions, which does not provide realistic vessel paths for image-based procedure guidance. Further, compared to the original X-ray images depicted in (a), the strong baseline network [17] still generates some
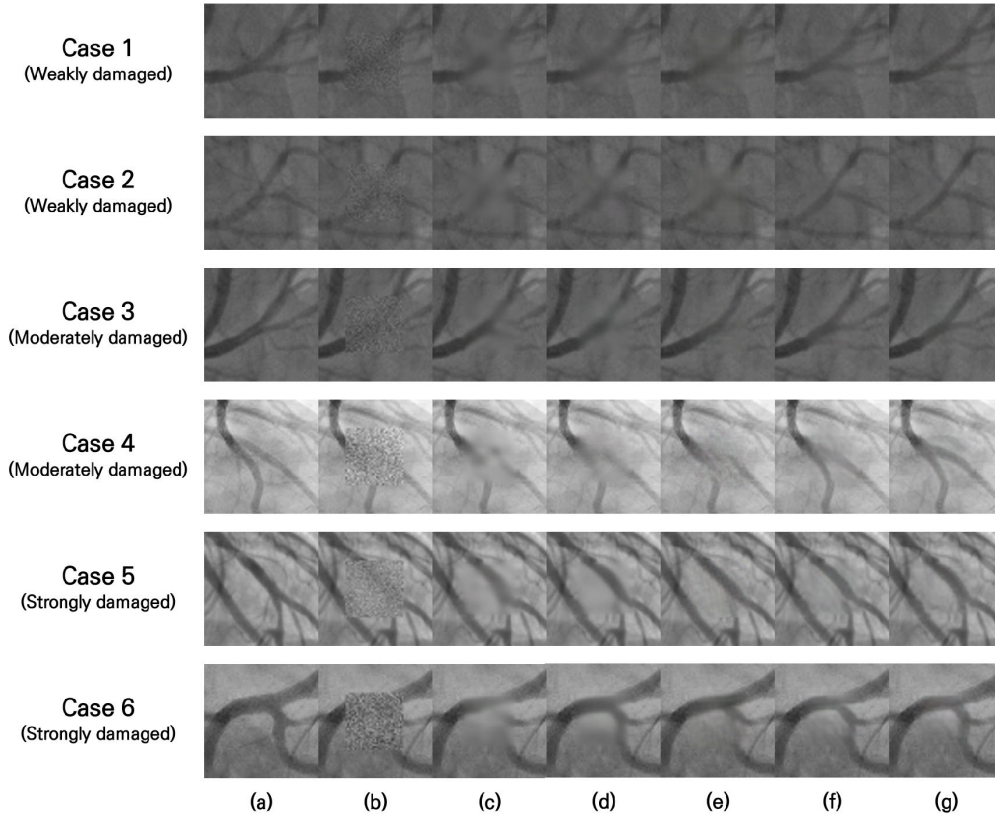
FIGURE 5. Qualitative results $\tilde{x}_i^{\text{damaged}}$ reconstructed using each baseline network using synthesized damaged images $x_i^{\text{patch-damaged}}$ under different settings: (a) intact X-ray images x_i^{patch} (GT); (b) synthesized images $x_i^{\text{patch-damaged}}$ (Input); (c) CE [15]; (d) IPA [18]; (e) HR [21]; (f) AOT [17]; (g) our proposed MAB network.

broken parts, and some of the vessels are narrower than the original ones because of the low intensity. On the other hand, the MAB network produces robust results without broken parts and does not generate narrow vessels.

D. EVALUATION IN THE ABSENCE OF VASCULAR INFORMATION

In this section, we consider the worst CAD case. To evaluate the robustness of the synthesis performance, the proposed method is used to reconstruct extreme cases featuring complete blockage of vessel information. We train our models to reconstruct blocked images, X_i^{blocked}. Here, the weight parameters of the various loss terms are taken to be $\lambda_1 = 1$, $\lambda_2 = 256$, $\lambda_3 = 0$, $\lambda_4 = 128$ and $\lambda_5 = 128$—these constants yield the best results.

As indicated by the quantitative results listed in Table 4, the proposed network achieves the best PNSR and MSE scores of 29.660 and 0.543, respectively, in the case of blocked images and damaged images, whereas its SSIM score of 0.873 is slightly lower than that of AOT (0.875) [17]. Moreover, even corresponding to completely blocked vessel information, all baseline networks exhibit the highest performance in terms of PSNR, MSE, and SSIM when trained in conjunction with vesselness-loss $\mathcal{L}_{\ell 1(V)}$. In particular, superior performance enhancements are effected on the baseline networks other

than MAB, when they are optimized with both $\mathcal{L}_{\ell 1(V)}$ and L_{SSIM}. In contrast, the MAB network performs the best when only $\mathcal{L}_{\ell 1(V)}$ is used, yielding PSNR, MSE, and SSIM scores of 29.850, 0.515, and 0.873, respectively. Its performance is still improved when both loss terms are used for training. Thus, for vessel-reconstruction tasks, the proposed vesselness-loss is demonstrated to be highly beneficial.

Additionally, the reconstructed results obtained from the networks are qualitatively compared in Fig. 6. From the reconstructed results of [15], [17], [18], and [21], we can observe that distortions remain in the edges or shapes of the vessels because the blocked images are reconstructed considering only the direction of blood flow in the vessels and contextual information of external regions. In contrast, the results of the MAB network are confirmed to contain almost no distortion in the vessel areas and to be reconstructed with similar intensity values as those of actual coronary arteries.

E. EVALUATION OF THE PROPOSED FULLY-AUTOMATIC RECONSTRUCTING METHODS

The proposed fully automated reconstruction process is evaluated in real CTO cases involving broken coronary arteries, $C = \{C^{\text{broken}} \cup C^{\text{vessel}}\}$. The process comprises three

TABLE 4. Quantitative comparison of various methods on the reconstruction of completely blocked images.

Methods	PSNR ↑	MSE ↓	SSIM ↑
CE [15]	28.809	0.630	0.855
CE + $\mathcal{L}_{SSIM}$	28.995 (+0.186)	0.610 (-0.020)	0.859 (+0.004)
CE + $\mathcal{L}_{\ell1(V)}$	29.120 (+0.311)	0.603 (-0.027)	0.858 (+0.003)
CE + $\mathcal{L}_{\ell1(V)}$ + L_{SSIM}	**29.148 (+0.339)**	**0.596 (-0.034)**	**0.862 (+0.007)**
IPA [18]	29.190	0.580	0.871
IPA + $\mathcal{L}_{SSIM}$	29.183 (-0.007)	0.590 (+0.010)	0.872 (+0.001)
IPA + $\mathcal{L}_{\ell1(V)}$	29.277 (+0.087)	0.570 (-0.010)	0.870 (-0.001)
IPA + $\mathcal{L}_{\ell1(V)}$ + L_{SSIM}	**29.365 (+0.175)**	**0.570 (-0.010)**	**0.872 (+0.001)**
HR [21]	29.251	0.600	0.864
HR + $\mathcal{L}_{SSIM}$	29.242 (-0.009)	0.592 (-0.008)	0.863 (-0.001)
HR + $\mathcal{L}_{\ell1(V)}$	29.421 (+0.170)	0.572 (-0.028)	0.864 (+0.000)
HR + $\mathcal{L}_{\ell1(V)}$ + L_{SSIM}	**29.457 (+0.205)**	**0.564 (-0.036)**	**0.866 (+0.002)**
AOT [17]	29.586	0.557	0.875
AOT + $\mathcal{L}_{SSIM}$	29.700 (+0.114)	0.526 (-0.031)	0.873 (-0.002)
AOT + $\mathcal{L}_{\ell1(V)}$	29.614 (+0.028)	0.546 (-0.011)	0.870 (-0.005)
AOT + $\mathcal{L}_{\ell1(V)}$ + L_{SSIM}	**29.738 (+0.152)**	**0.538 (-0.019)**	**0.878 (+0.003)**
MAB	29.660	0.543	0.873
MAB + $\mathcal{L}_{SSIM}$	29.753 (+0.093)	0.519 (-0.015)	0.872 (-0.001)
MAB + $\mathcal{L}_{\ell1(V)}$	**29.850 (+0.190)**	**0.515 (-0.019)**	**0.873 (+0.000)**
MAB + $\mathcal{L}_{\ell1(V)}$ + L_{SSIM}	29.767 (+0.107)	0.519 (-0.015)	0.872 (-0.001)

Notes: Even when vessel information is completely blocked, our novel MAB networks achieve superior performance compared to those of other baseline networks (CE[15], IPA[18], HR[21] and AOT[17]). Furthermore, optimizing the vesselness-loss function leads to performance improvements in all networks.
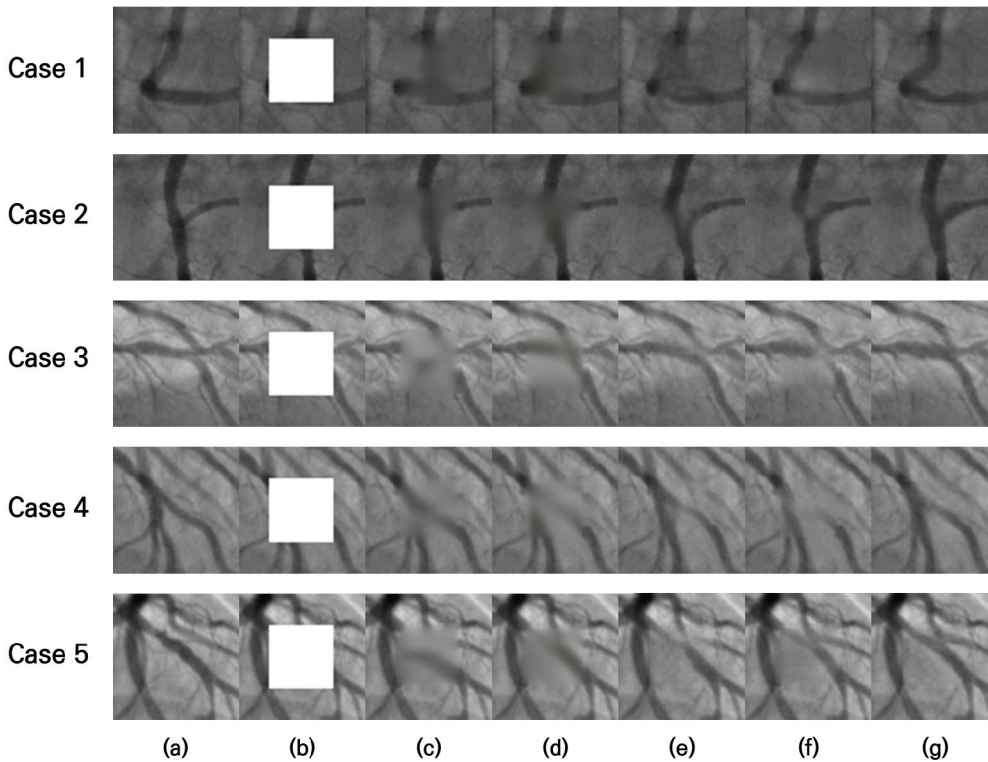

(a) (b) (c) (d) (e) (f) (g)

FIGURE 6. Qualitative results $\tilde{x}_i^{blocked}$ of baseline networks using completely blocked images $X_i^{blocked}$: (a) ideal patches x_i^{patch} containing intact vessel information; (b) blocked patches $x_i^{blocked}$ without vessel information; (c) CE [15]; (d) IPA [18]; (e) HR [21]; (f) AOT [17]; (g) proposed MAB network.

steps. First, we utilize our previous studies on ROI detection methods. Thus, tip detection sub-methods are used to search for broken parts. Second, target areas $X_i^{blocked}$ are defined where the broken parts are to be reconstructed. Finally, those target areas $X_i^{blocked}$ are reconstructed realistically to be similar to the broken coronary arteries, $C = \{C^{broken} \cup$
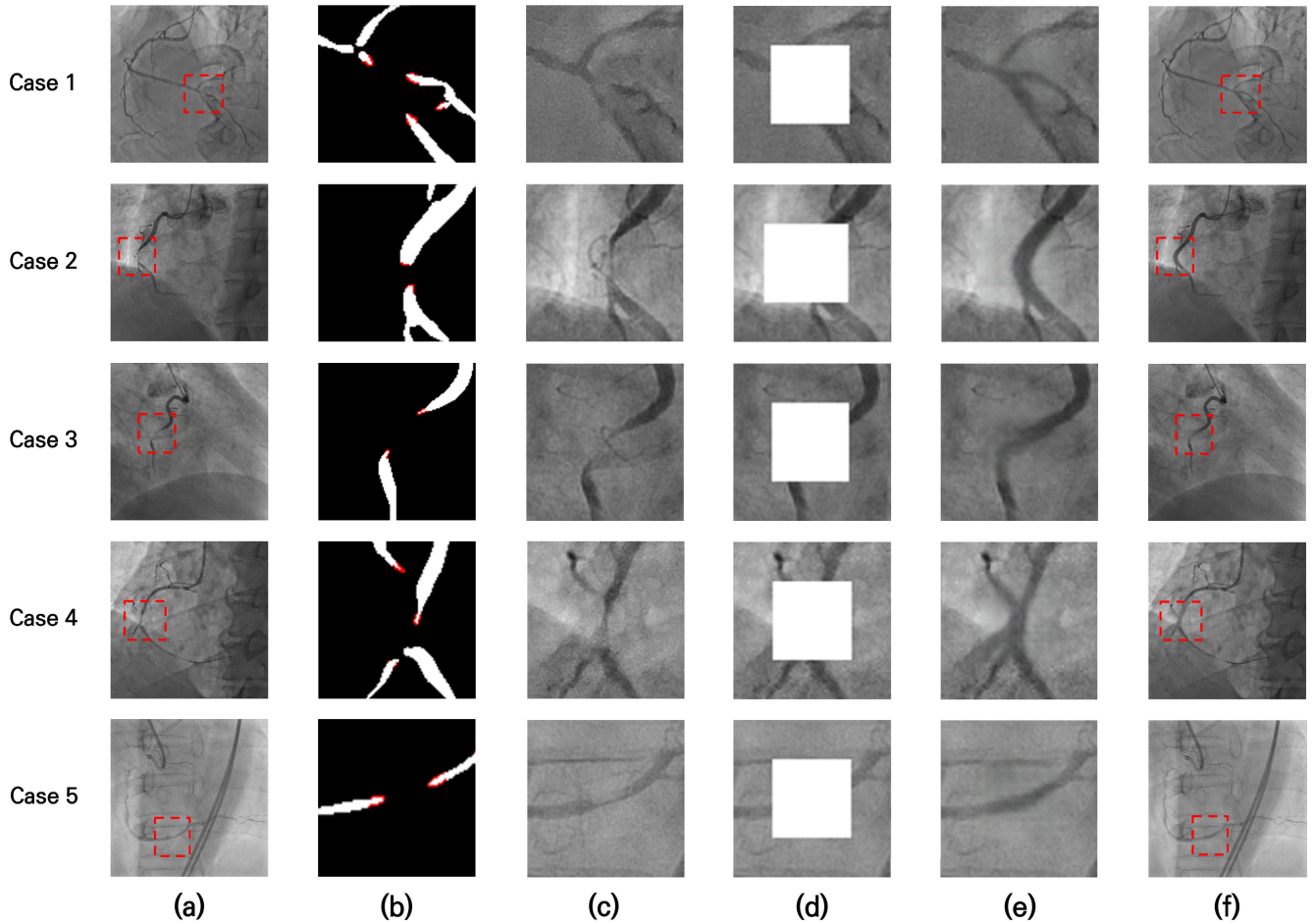
FIGURE 7. Results of the proposed fully automated reconstruction process in X-ray images: (a) 2D X-ray images that contain broken coronary arteries $C = \{C^{\text{broken}} \cup C^{\text{vessel}}\}$; (b) automatically detected ROI areas via ROI detection methods, with tip points colored in red; (c) corresponding X-ray patches where broken parts exist; (d) target areas X_i^{blocked} where a block is blended with broken parts; (e) realistically reconstructed outputs $\tilde{X}_i^{\text{blocked}}$, showing vascular path; (f) reconstructed 2D X-ray images containing restored coronary arteries indistinguishable from ideal $C = \{C^{\text{vessel}}\}$.

$C^{\text{vessel}}\}$, relative to the ideal case, $C = \{C^{\text{vessel}}\}$. We block the target areas to assume no meaningful vessel information present on the broken parts.

The reconstructed results of the proposed MAB network are presented in Fig. 7. First, the broken parts in the X-ray image are detected precisely via ROI detection, including the core broken coronary arteries. Next, the vascular paths are very realistically reconstructed even when the target areas are completely blocked, owing to the explicit guidance provided by our vesselness-loss.

F. ABLATION STUDIES

In this section, we conduct ablation studies to verify and demonstrate the effectiveness of our vesselness-loss function. The previous quantitative comparison in Table 3 and 4 demonstrates that using the vesselness-loss function, all models demonstrate improved performance in terms of PSNR, MSE, and SSIM when the vessel information is damaged or completely blocked. It is also observed that, when vesselness-loss is used in conjunction with SSIM loss, some models exhibit larger performance improvements. Even if

model performance does not improve when trained with both SSIM loss and vesselness-loss, it still improves when trained using only vesselness-loss.

We also evaluate whether the vesselness-loss constrains the models to foster a vessel-specific reconstruction capability as follows. First, we configure the other factors, i.e., internal parameters of the model architecture, remaining loss functions, and hyper-parameters, such as learning rate and mini-batch size, to be identical. Then, we train the network under four configurations: 1) without both $\mathcal{L}_{\ell1(V)}$ and L_{SSIM}, 2) with only L_{SSIM}, 3) with only $\mathcal{L}_{\ell1(V)}$, and 4) with both $\mathcal{L}_{\ell1(V)}$ and L_{SSIM}.

Figs. 8(a) and (b) present an X-ray image and the synthesized blocked images. The model trained without $\mathcal{L}_{\ell1(V)}$ and L_{SSIM} generates blurry results, as depicted in Fig. 8(c). When trained using only L_{SSIM}, the model yields visually improved results, as depicted in Fig. 8(d), compared to those depicted in (c). On the other hand, for the models trained using $\mathcal{L}_{\ell1(V)}$, the edges of the vessel regions are clearly reconstructed, as shown in Fig. 8(e), compared to those depicted in (c) and even (d). Moreover, the intensity
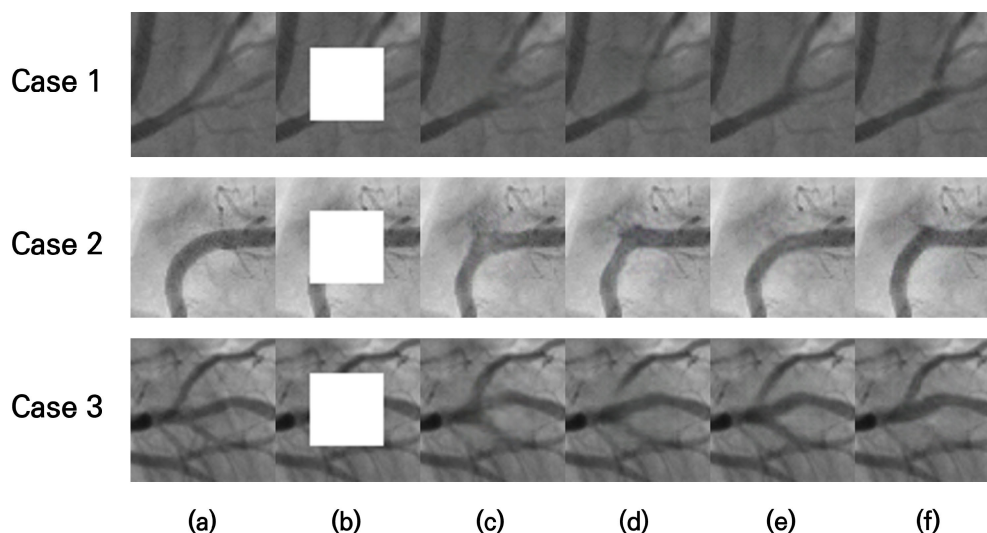
FIGURE 8. Multiple results from ablation studies to visualize effectiveness of our novel vesselness-loss function on our novel MAB network: (a) X-ray images X_i^{patch}; (b) blocked images $X_i^{blocked}$; (c) blurry results of model trained without both f the model trained without both; (d) blurry results for where broken parts exist of model trained only with L_{SSIM}; (e) clear results containing realistically restored intensity values of model trained only with $\mathcal{L}_{\ell1(V)}$; (f) results with sharply represented edges in vessel regions of model trained with both $\mathcal{L}_{\ell1(V)}$ and L_{SSIM}.

values in vessel regions are reconstructed to be identical to those of (a) without broken regions. This is because vesselness-loss focuses on recovering the vessel structure and intensity in the vessel region by minimizing the difference in the vesselness probability map between the input image and model output. To describe further, SSIM loss may train models to learn luminance and even noise, besides structure, in the X-ray images. Meanwhile, the models can focus on vessel-specific structure when optimized using vesselness-loss, because it minimizes the pixel-wise difference only within the vessels. Finally, in cases 1 and 3, it is verified that training using $\mathcal{L}_{\ell1(V)}$ and L_{SSIM} are highly synergetic in vessel reconstruction, as revealed by the comparison between Fig. 8(f) and (d, e)—the former depicts the results obtained using the model trained using both $\mathcal{L}_{\ell1(V)}$ and L_{SSIM}, whereas the latter depicts those obtained using the model trained using either $\mathcal{L}_{\ell1(V)}$ or L_{SSIM}.

V. DISCUSSION

Partially invisible coronary arteries in X-ray images have been reported to be one of the main challenges that hinder successful image-guided PCI procedures. The diagnosis of coronary arteries is challenging for two reasons. Firstly, cardiac motion artifacts or intrinsic limitations of X-ray images, e.g., noise, adversely affect the quality of X-ray images, thereby leading to indistinct vessel regions. Secondly, in patients with severe CAD, such as CTO, the vessel regions are completely blocked, preventing the visualization of the vessel regions via the injection of contrastive substances into the regions.

To address the aforementioned issues, we propose a method that uses a GAN-based approach to reconstruct blocked coronary artery regions directly using only single X-ray modalities. The proposed method aims to reconstruct blocked regions into ideal regions automatically, or depict broken coronary arteries as ideal coronary arteries. For very realistic reconstruction and clinical aid, we propose an MAB network that accounts for both global and local features, e.g., direction of the vessels and context information, respectively. Additionally, a novel vesselness-loss function is proposed to induce the networks to focus on sparse vessel regions while learning reconstruction and, specifically, to generate synthetic vessels with similar intensity values or shapes to those of real vessels.

Via quantitative and qualitative experiments, it is confirmed that the proposed MAB networks, which encodes multi-scale images for reconstruction tasks, exhibits the best performance even in cases of damaged or blocked vessel information. Furthermore, it is verified that our novel vesselness-loss function, which is designed to focus on the vessel regions, is effective, exhibiting distinctly improved performance in quantitative evaluation. We further demonstrate that vesselness-loss can control the model during the vessel reconstruction process—when the vesselness-loss is used, the reconstructed outputs recover similar intensities or shapes as those of the real vessels. Further, vesselness-loss generates synergistic results when optimized in conjunction with other objective functions, e.g., SSIM loss. However, SSIM loss conserves the structures of entire images without taking consideration into vessel-structures—thus, during reconstruction, noise or inherent luminance of X-ray images inevitably are included during the reconstruction process. On the other hand, as vesselness-loss conserves the structures only in vessel regions, the reconstructed results conserve robust vessel edges. In our experiments, when the weighted constant λ_5 of the vesselness-loss is larger than the weighted

constants λ_3 of the SSIM loss, the vessels are reconstructed realistically.

The proposed fully automatic method reconstructs broken parts of the coronary artery by integrating our previously proposed ROI detection method and a novel GAN-based model. Excluding input-output under our experiment environment, only 0.27 s (3.7 Hz) is required per image to perform the entire process. Most of the operating time is attributed to tip point detection, which identifies broken parts. Notably, compared to the times required by conventional registration-based methods, that of the proposed methods is acceptable for real clinical situations.

However, the proposed method suffers from certain limitations. We have not considered cases in which the vessel is blocked from the initial point of the catheter. In such cases, the broken parts can be lengthy, which is not addressed by the current method. In addition, extremely rare cases, involving multiple broken parts, including CTO and stumpless cases, are not considered. Meanwhile, as mentioned previously, detection of broken parts via tip detection is a major bottleneck in terms of operating time. To overcome these limitations, we intend to construct a pipeline consisting of two networks—a module that detects multiple broken parts in multiple sites over an entire 2D X-ray image, and a deep and versatile inpainting network that reconstructs various broken parts, including lengthy vessels in future works.

In summary, the current paper is important from both engineering and clinical perspectives. From the engineering perspective, an MAB network and vesselness-loss are proposed that encode both vessel-specific regions and contextual information and guide networks to focus on vessel-specific areas to reconstruct realistic X-ray images. From the clinical perspective, we propose the first fully automatic framework that receives only a single X-ray image to reconstruct broken coronary artery regions. In addition, the proposed method requires the smallest amount of time, 0.27 s (3.7 Hz), among all existing alternatives, but reconstructs the most robust and realistic vascular paths, even in the worst CTO cases. We expect the proposed method to be utilized for image-guided procedure or diagnosis systems in real clinical sites.

VI. CONCLUSION

X-ray images are the primary standard images used to guide PCI procedures during the treatment of blocked coronary arteries. Unfortunately, it is challenging to detect and identify of accurate vascular paths for damaged or invisible vessels during the procedure due to inherent limitations of X-ray images. There have been proposed several registration-based techniques that match the features of 2D X-ray images with those of 3D CT images, however, they are difficult to apply in real-world emergency situations because registration is time-consuming and requires both modalities.

For convenience and simplicity, we propose a fully-automatic GAN-based reconstruction pipeline that exploits only single modality, 2D X-ray images. First, we devise an MAB network, which robustly and realistically reconstructs broken coronary arteries as ideal ones. Then, we propose a novel vesselness-loss that compels the model to focus on sparse vessels. Under an extensive experiments using in-house 2D X-ray angiography vessel data, we confirm that our both proposed methods contribute to outstanding vessel reconstruction performance both quantitatively and qualitatively, even in the worst situation where vessel information is completely blocked. The fully automatic pipeline, from broken region detection to reconstruction, is achieved when the proposed methods are adopted to our previous tip point detection module.

In conclusion, the proposed method is highly worth to be considered to have a significant clinical value as it can be operated in a fully automatic fashion to reconstruct very realistic vessel information, which is sufficient to support an image-guided procedure. In the future, we expect our proposed methods to be applied to real clinical sites for image-based procedure guidance.

REFERENCES

[1] N. Townsend, L. Wilson, P. Bhatnagar, K. Wickramasinghe, M. Rayner, and M. Nichols, "Cardiovascular disease in Europe: Epidemiological update 2016," *Eur. Heart J.*, vol. 37, no. 42, pp. 3232–3245, 2016.

[2] F. Sanchis-Gomar, C. Perez-Quilis, R. Leischik, and A. Lucia, "Epidemiology of coronary heart disease and acute coronary syndrome," *Ann. Transl. Med.*, vol. 4, no. 13, p. 256, Jul. 2016.

[3] G. A. Mensah, G. A. Roth, and V. Fuster, "The global burden of cardiovascular diseases and risk factors: 2020 and beyond," *J. Amer. College Cardiol.*, vol. 74, no. 20, pp. 2529–2532, Nov. 2019.

[4] N. Baka, C. T. Metz, C. Schultz, L. Neefjes, R. J. van Geuns, B. P. F. Lelieveldt, W. J. Niessen, T. van Walsum, and M. de Bruijne, "Statistical coronary motion models for 2D+t/3D registration of X-ray coronary angiography and CTA," *Med. Image Anal.*, vol. 17, no. 6, pp. 698–709, Aug. 2013.

[5] N. Baka, B. P. F. Lelieveldt, C. Schultz, W. Niessen, and T. van Walsum, "Respiratory motion estimation in X-ray angiography for improved guidance during coronary interventions," *Phys. Med. Biol.*, vol. 60, no. 9, pp. 3617–3637, May 2015.

[6] U. Mitrović, Z. Špiclin, B. Likar, and F. Pernuš, "3D-2D registration of cerebral angiograms: A method and evaluation on clinical images," *IEEE Trans. Med. Imag.*, vol. 32, no. 8, pp. 1550–1563, Aug. 2013.

[7] J. Zhu, H. Li, D. Ai, Q. Yang, J. Fan, Y. Huang, H. Song, Y. Han, and J. Yang, "Iterative closest graph matching for non-rigid 3D/2D coronary arteries registration," *Comput. Methods Programs Biomed.*, vol. 199, Feb. 2021, Art. no. 105901.

[8] T. Park, S. Khang, H. Jeong, K. Koo, J. Lee, J. Shin, and H. C. Kang, "Deep learning segmentation in 2D X-ray images and non-rigid registration in multi-modality images of coronary arteries," *Diagnostics*, vol. 12, no. 4, p. 778, Mar. 2022.

[9] D. Jeong, D. Kim, J. Ryu, and K. H. Cho, "Deep-learning-based registration of diagnostic angiogram and live fluoroscopy for percutaneous coronary intervention," *IEEE Access*, vol. 9, pp. 103465–103480, 2021.

[10] W. Wu, J. Zhang, W. Peng, H. Xie, S. Zhang, and L. Gu, "CAR-Net: A deep learning-based deformation model for 3D/2D coronary artery registration," *IEEE Trans. Med. Imag.*, vol. 41, no. 10, pp. 2715–2727, Oct. 2022.

[11] K. Han, J. Jeon, Y. Jang, S. Jung, S. Kim, H. Shim, B. Jeon, and H.-J. Chang, "Reconnection of fragmented parts of coronary arteries using local geometric features in X-ray angiography images," *Comput. Biol. Med.*, vol. 141, Feb. 2022, Art. no. 105099.

[12] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 1–9.

[13] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 4396–4405.

[14] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2242–2251.

[15] D. Pathak, P. Krähenbühl, 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.

[16] K. Nazeri, E. Ng, T. Joseph, F. Z. Qureshi, and M. Ebrahimi, "EdgeConnect: Generative image inpainting with adversarial edge learning," 2019, *arXiv:1901.00212*.

[17] Y. Zeng, J. Fu, H. Chao, and B. Guo, "Aggregated contextual transformations for high-resolution image inpainting," *IEEE Trans. Vis. Comput. Graphics*, vol. 29, no. 7, pp. 3266–3280, Jul. 2023.

[18] K. Armanious, V. Kumar, S. Abdulatif, T. Hepp, S. Gatidis, and B. Yang, "IpA-MedGAN: Inpainting of arbitrary regions in medical imaging," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2020, pp. 3005–3009.

[19] B. Sim, G. Oh, J. Kim, C. Jung, and J. C. Ye, "Optimal transport driven CycleGAN for unsupervised learning in inverse problems," *SIAM J. Imag. Sci.*, vol. 13, no. 4, pp. 2281–2306, Jan. 2020.

[20] Q. Wang, Y. Chen, N. Zhang, and Y. Gu, "Medical image inpainting with edge and structure priors," *Measurement*, vol. 185, Nov. 2021, Art. no. 110027.

[21] 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. 7505–7514.

[22] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proc. 27th Annu. Conf. Comput. Graph. Interact. Techn.*, 2000, pp. 417–424.

[23] A. A. Efros and W. T. Freeman, "Image quilting for texture synthesis and transfer," in *Proc. 28th Annu. Conf. Comput. Graph. Interact. Techn.*, Aug. 2001, pp. 341–346.

[24] C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman, "PatchMatch: A randomized correspondence algorithm for structural image editing," *ACM Trans. Graph.*, vol. 28, no. 3, p. 24, 2009.

[25] J. Hays and A. A. Efros, "Scene completion using millions of photographs," *Commun. ACM*, vol. 51, no. 10, pp. 87–94, Oct. 2008.

[26] S. Darabi, E. Shechtman, C. Barnes, D. B. Goldman, and P. Sen, "Image melding: Combining inconsistent images using patch-based synthesis," *ACM Trans. Graph.*, vol. 31, no. 4, pp. 1–10, Aug. 2012.

[27] 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.

[28] 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.

[29] H. Liu, B. Jiang, Y. Xiao, and C. Yang, "Coherent semantic attention for image inpainting," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 4169–4178.

[30] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. Huang, "Free-form image inpainting with gated convolution," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 4470–4479.

[31] Y. Ma, X. Liu, S. Bai, L. Wang, D. He, and A. Liu, "Coarse-to-fine image inpainting via region-wise convolutions and non-local correlation," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, Aug. 2019, pp. 3123–3129.

[32] X. Guo, H. Yang, and D. Huang, "Image inpainting via conditional texture and structure dual generation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 14114–14123.

[33] Z. Feng, S. Chi, J. Yin, D. Zhao, and X. Liu, "A variational approach to medical image inpainting based on Mumford-Shah model," in *Proc. Int. Conf. Service Syst. Service Manage.*, Jun. 2007, pp. 1–5.

[34] N. Guizard, K. Nakamura, P. Coupé, V. S. Fonov, D. L. Arnold, and D. L. Collins, "Non-local means inpainting of MS lesions in longitudinal image processing," *Frontiers Neurosci.*, vol. 9, p. 456, Dec. 2015.

[35] M. Arnold, A. Ghosh, S. Ameling, and G. Lacey, "Automatic segmentation and inpainting of specular highlights for endoscopic imaging," *EURASIP J. Image Video Process.*, vol. 2010, Dec. 2010, Art. no. 814319.

[36] K. Armanious, Y. Mecky, S. Gatidis, and B. Yang, "Adversarial inpainting of medical image modalities," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 3267–3271.

[37] J. Li, F. Fang, K. Mei, and G. Zhang, "Multi-scale residual network for image super-resolution," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 517–532.

[38] J. Qin, Y. Huang, and W. Wen, "Multi-scale feature fusion residual network for single image super-resolution," *Neurocomputing*, vol. 379, pp. 334–342, Feb. 2020.

[39] Z. Hui, J. Li, X. Wang, and X. Gao, "Image fine-grained inpainting," 2020, *arXiv:2002.02609*.

[40] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent*. Cambridge, MA, USA: Springer, Oct. 1998, pp. 130–137.

[41] Y. Sato, S. Nakajima, H. Atsumi, T. Koller, G. Gerig, S. Yoshida, and R. Kikinis, "3D multi-scale line filter for segmentation and visualization of curvilinear structures in medical images," in *Proc. CVRMed-MRCAS*. Grenoble, France: Springer, 1997, pp. 213–222.

[42] J. Yang, M. Huang, J. Fu, C. Lou, and C. Feng, "Frangi based multi-scale level sets for retinal vascular segmentation," *Comput. Methods Programs Biomed.*, vol. 197, Dec. 2020, Art. no. 105752.

[43] K. B. Khan, A. A. Khaliq, and M. Shahid, "A novel fast GLM approach for retinal vascular segmentation and denoising," *J. Inf. Sci. Eng.*, vol. 33, no. 6, pp. 1611–1627, 2017.

[44] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[45] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5987–5995.

[46] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 11966–11976.

[47] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.

[48] A. Veit, M. J. Wilber, and S. Belongie, "Residual networks behave like ensembles of relatively shallow networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 1–9.

[49] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 1125–1134.

[50] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida, "Spectral normalization for generative adversarial networks," 2018, *arXiv:1802.05957*.

[51] J. H. Lim and J. C. Ye, "Geometric GAN," 2017, *arXiv:1705.02894*.

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

[53] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and S. Wenzhe, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 4681–4690.

[54] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[55] M. Abdullah-Al-Mamun, V. Tyagi, and H. Zhao, "A new full-reference image quality metric for motion blur profile characterization," *IEEE Access*, vol. 9, pp. 156361–156371, 2021.

KYUNGHOON HAN (Member, IEEE) received the B.S. degree in digital information engineering from the Hankuk University of Foreign Studies, South Korea, in 2013. He is currently pursuing the Ph.D. degree in medical science with Yonsei University, Seoul, South Korea. Since 2018, he has been a Researcher with the Medical Image Computing and Analysis Laboratory, CONNECT-AI Research Center, Yonsei University College of Medicine. His research interests include the application of computer vision techniques and deep learning from image processing, geometric analysis, and segmentation tasks in medical images.

HEEJOON KOO received the M.Eng. degree in electronic and electrical engineering from University College London (UCL), U.K., in 2021. Then, he joined the CONNECT-AI Research Center, Yonsei University College of Medicine, South Korea. His research interest includes the overall field of deep learning and domains where deep learning is applicable and even revolutionize.

SUNGHEE JUNG (Member, IEEE) received the B.S. degree in computer science engineering from the Hankuk University of Foreign Studies, South Korea, in 2013, and the Ph.D. degree in medical science from Yonsei University, South Korea, in 2020. She was a Researcher with Ontact Health Inc., South Korea, and a Visiting Professor with the CONNECT-AI Research Center, Yonsei University. She is currently a Professor with the School of Electrical Engineering, Kookmin University, Seoul, South Korea. Her research interests include the application of computer vision techniques from image processing and deep learning to noise reduction and motion correction in medical images.
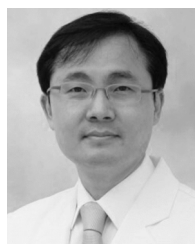
HYUNG-BOK PARK received the M.D. and M.M. degrees from the College of Medicine, Catholic Kwandong University, Gangneung, South Korea, in 2006 and 2012, respectively, and the Ph.D. degree in medicine from the Graduate School of Medicine, Yonsei University, Seoul, South Korea, in 2023. From 2012 to 2013, he was a Cardiac Imaging Research Fellow with the Cedars-Sinai Hospital, Los Angeles, CA, USA. From 2014 to 2017, he was an Assistant Professor with the Department of Cardiology, Myongji Hospital. Since 2017, he has been an Associate Professor with the Cardiology Department, Catholic Kwandong University International St. Mary's Hospital, Incheon, South Korea. His research interests include cardiovascular imaging (coronary CTA, cardiac MR, and echocardiography), interventional cardiology, and medical image processing.

YOUNGTAEK HONG (Member, IEEE) received the B.S. degree in digital information engineering from the Hankuk University of Foreign Studies, Gyeonggi-do, Republic of Korea, in 2012, and the Ph.D. degree in medical science from the Yonsei University College of Medicine, Seoul, Republic of Korea, in 2018. He was a Visiting Scholar with the Biomedical Imaging Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA, from 2017 to 2018. In 2018, he was a Postdoctoral Research Fellow with the Yonsei University College of Medicine, where he has been a Research Assistant Professor, since 2020. His research interests include computer vision, medical image analysis, and radiomics.

HACKJOON SHIM received the B.S. and Ph.D. degrees from Seoul National University, Seoul, South Korea, in 1993 and 2007, respectively. From 2008 to 2010, he was a Research Professor with the Department of Electrical and Computer Engineering, College of Engineering, Seoul National University. Since 2011, he has been a Research Professor with the Severance Hospital, Yonsei University, Seoul. His research interests include computer vision and medical image analysis.

BYUNGHWAN JEON (Member, IEEE) received the B.S. degree in computer science engineering from the Hankuk University of Foreign Studies, South Korea, in 2013, and the Ph.D. degree in medical science from Yonsei University, Seoul, South Korea, in 2019. From 2012 to 2019, he was a Researcher with the Medical Image Computing and Analysis Laboratory, CONNECT-AI Research Center, Seoul. From 2020 to 2021, he was an Assistant Professor with the Department of Computer Science, Kyungil University, Gyeongsan, South Korea. Since 2019, he has been a Postdoctoral Researcher with the Medical Image Computing and Analysis Laboratory, CONNECT-AI Research Center. Since 2022, he has been an Assistant Professor with the Division of Computer Engineering, Hankuk University of Foreign Studies, Yongin, South Korea. His research interests include the application of computer vision techniques, from Bayesian estimation, geometric analysis, and deep neural networks to localization, tracking, and segmentation tasks in medical images.

HYUK-JAE CHANG received the M.D. and M.M. degrees from the College of Medicine, Yonsei University, Seoul, South Korea, in 1994 and 1999, respectively, and the Ph.D. degree in medicine from the Graduate School of Medicine, Ajou University, Suwon, South Korea, in 2003. From 2002 to 2004, he was an Assistant Professor with the Department of Internal Medicine, College of Medicine, Ajou University. From 2004 to 2009, he was an Assistant Professor with the Department of Internal Medicine, Seoul National University College of Medicine, Seoul. From 2007 to 2008, he was a Postdoctoral Fellow with the Department of Internal Medicine, Johns Hopkins University, Baltimore, MD, USA. Since 2009, he has been a Professor with the Department of Internal Medicine, Yonsei University College of Medicine. Since 2016, he has been the CIO of the Division of Medical Information and Technology, Severance Hospital. From 2017 to 2020, he was the CEO of PHI Digital Healthcare Inc., Seoul. He is currently the CEO and the Founder of Ontact healthcare. His research interests include cardiovascular disease, pulmonary hypertension, and medical image processing.

• • •