Deep-Learning based Restoration Methods for Missing Vessels in 2D X-ray Angiography Images

Kyunghoon Han

Department of Medical Science

The Graduate School, Yonsei University

Deep-Learning based Restoration Methods for Missing Vessels in 2D X-ray Angiography Images

Directed by Professor Hyuk-Jae Chang

The Doctoral Dissertation

submitted to the Department of Medical Science,

the Graduate School of Yonsei University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Medical Science

Kyunghoon Han

December 2023

This certifies that the Doctoral Dissertation of Kyunghoon Han is approved.

Thesis Supervisor: Hyuk-Jae Chang

Thesis Committee Member#1: In Hyun Jung

Thesis Committee Member#2: Dosik Hwang

Thesis Committee Member#3: Hackjoon Shim

Thesis Committee Member#4: Byunghwan Jeon

The Graduate School

Yonsei University

December 2023

ACKNOWLEDGMENTS

돌이켜 생각해보면, 저는 참 부족한 사람이였던 것 같습니다. 아무것도 모르는 철없는 학부생 인턴 연구원이었던 저에게는 학부 시절 배운 지식이 제가 아는 전부였고, 제가 보고 듣는 것들이 전부인 줄 알았던 한없이 작은 사람이였던 것 같습니다.

그러나 제가 좋은 환경에서 알맞은 방향으로 나아갈 수 있도록 이끌어주시며 힘들 때마다 저의 편에서 저를 헤아려 주신 지도교수님이신 장혁재 교수님과, 컴퓨터 비전을 기반으로 공학 연구를 할 수 있도록 지도해주신 심학준 교수님, 어려운 일이 생길 때마다 제 일을 본인 일처럼 고민해주시고 도움을 주신 사수인 병환이형, 서로 지식을 공유하며 연구를 한 선후배 연구원들을 만남으로써 부족한 제가 학위 과정을 무사히 마칠 수 있는 것을 보면, 한편으로는 저는 운이 좋은 사람이였던 것 같습니다.

제 인생에서 가장 중요한 시기였던 박사학위과정동안 인연이라는 계기를 통해 함께 연구를 진행한 연구실 멤버들 (성민이형, 영택이형, 성희누나, 영걸이형, 세근이형, 지나, 현석, 가은, 주영, 다운, 자영, 지연, 시현, 희준, 재익)과 큰 프로젝트를 함께 진행한 선생님들 (김태영 선생님, 김지영 선생님 등) 및 행적적으로 도움을 주시는 행정팀 선생님들 (이영은 선생님, 안진영 선생님 등)과 제가 힘들 때마다 바쁜 시간을 내서 제 옆에서 저를 응원해주고 격려해준 저의 친구들 (요한, 준명, 제구, 경서, 종석 등)까지 모든 분들께 진심으로 감사의 인사를 드립니다.

또한, 공학밖에 모르던 제가 공학과 의학의 융합 연구를 통해 사람을 향하는 연구를 진행할 수 있도록 임상 연구의 관점에서 많은 조언을 주신 박형복 교수님과, 정인현 교수님, 허란 교수님 및 논문 심사과정에서 아낌없는 조언을 해주신 황도식 교수님께도 감사의 말씀을 드립니다.

처음 대학원 생활을 시작할 때, 참을성이 부족하였던 스스로에게 "마부작침의 자세로 연구를 하겠노라." 다짐하며 열심히 연구에 매진했던 저에게 있어 대학원에서의 생활은 학술적 성숙과 더불어, 참을성을 가르쳐 주었으며, 정서적, 인격적으로도 많이 가르침을 준 시간이였던 것 같습니다.

독주보다는 협주가 아름다운 멜로디를 낼 수 있다는 가르침을 주신 장혁재 교수님의 말씀을 토대로, 독주보다는 협주라는 수식어가 어울리는 연구자가 될 수 있도록, 그리고 박사학위과정동안 성숙해진 자세로 저희 연구실을 빛낼 수 있도록 끊임없이 노력하는 겸손한 연구자가 되도록 노력하겠습니다.

마지막으로, 아직도 한없이 철부지 아들이며 매일 장난만 치는 오빠이자 형인 저를 이 자리에 있게 해주신 제가 정말 사랑하는 아버지, 어머니, 동생, 그리고 땡큐에게 진심을 다해 감사의 말씀을 드립니다. 저는 제가 아버지, 어머니의 아들이라는 것이 그 누구보다 자랑스럽습니다. 남들과는 조금 다른 학창시절을 보냈음에도 불구하고, 제가 하고자 하는 일들을 응원해주시고 지원해주시는 부모님과 묵묵히 저를 도와준 동생들에게 자랑스러운 아들이자, 오빠이자, 형이 될 수 있도록 노력하겠습니다. 그리고 제 가장 친한 친구이며 즐거울 때나 슬플 때나 묵묵히 제 바로옆에서 저를 바라봐주는 저의 강아지 땡큐에게도 진심으로 고맙다는 인사를 전합니다.

한경훈 올림

<Table of Contents>

List of Figures

List of Tables

ABSTRACT

Deep-Learning based Restoration Methods for Missing Vessels in 2D X-ray Angiography Images

Kyunghoon Han

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Hyuk-Jae Chang)

Purpose - Coronary artery procedures are primarily performed based on 2-dimensional (2D) X-ray angiography images. However, inherent noise and cardiac motion artifacts in these images often make it challenging to operate image-guided catheter procedures. In specific, in the case of chronic total occlusion (CTO), in which the coronary artery is completely occluded, the contrast does not pass and the path of the vessel is indistinguishable, therefore vessel perforation may occur frequently due to the excessive wiring manipulation. To address these challenges, this paper introduces a deep learning-based frameworks designed to restore missing coronary arteries by reconnecting the fragmented coronary artery segmentation results and to reconstruct the 2D X-ray angiography images to facilitate more precise image-guided interventions.

Methods - First, this paper proposes a method to restore the missing coronary arteries by reconnecting the broken segmentation results. This is clinically useful as it provides a more accurate coronary artery area and vascular path to operators. While a simple convolutional neural network (CNN) based segmentation model can be obtained through supervised learning leveraging a pair of images and their respective ground truths, it faces difficulties in inferring the intricate and thin structures of coronary arteries, often yielding partially fragmented results. In addition, the coronary arteries, which should be connected smoothly and seamlessly, are partially invisible since vessel information does not exist in the image due to lesions such as CTO. To this end, this research proposes sophisticated post-processing techniques that use local details and geometric priors to reconnect partially broken coronary arteries.

Secondly, a generative adversarial networks (GAN) based 2D X-ray reconstruction model is proposed. It restores vascular regions in which it is indistinct due to image noise or lesions, and its clinical utility lies in its potential to provide a clearer vessel path for operators. Unlike the previous method, this reconstruction model is learned via unsupervised learning, solely utilizing X-ray images. This method aims to restore invisible vessel regions using only X-ray images, without necessitating any other external modalities.

Results - The proposed post-processing methods for reconnecting fragmented coronary arteries showed promising results in performance evaluation. Notably, the number of disconnected regions significantly reduced from 2.308 per image in the initial segmentation to 1.197 after applying the proposed method. Furthermore, when the post-processing methods was applied to the initial segmentation results of state-of-the-art segmentation models, the dice similarity coefficient (DSC) improved by 2.33 times in average, and the Jaccard index (JI) increased by an average of 2.88 times in the disconnected regions.

Moreover, the proposed reconstruction model demonstrated its superiority compared to other state-of-the-art reconstruction models in three image quality metrics such as peak signal-to-noise ratio (PSNR), mean squared error (MSE) and similarity indexing methods (SSIM). Qualitatively, it was observed that the proposed method is capable of preserving the vessel centerline which is clinically crucial element for the interventional procedure guidance.

Conclusions - Our proposed two methods provides vascular segmentation results or X-ray images that are clearly valuable for clinicians during procedures. Therefore, we expect that these techniques will be applied in real clinical settings in the near future.

Key words : deep learning, 2d x-ray angiography, coronary artery, chronic total occlusion, cardiovascular disease

Deep-Learning based Restoration Methods for Missing Vessels in 2D X-ray Angiography Images

Kyunghoon Han

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Hyuk-Jae Chang)

I. INTRODUCTION

Coronary artery diseases (CADs) is the causes of death in the world [1]–[3]. CAD occurs when cholesterol, fats, or various types of plaques [4] accumulate in blood vessels. If it becomes severe, it can obstruct blood flow and lead to vessel blockage, eventually resulting in chronic total occlusion (CTO). To treat CAD, a percutaneous coronary intervention (PCI) is performed improve blood flow by inserting a catheter into the blood vessel for its expansion. During procedures, 2-dimensional (2D) X-ray angiography images serve as a gold standard for localizing catheter tip and figuring out dynamically moving coronary artery structure in real time. The entire process of procedures are shown in Fig. 1.

However, due to the projection of the 3D cardiac anatomy onto a two-dimensional plane, there is an inevitable loss in spatial information. This causes significant challenges to the interpretation of the originally 3D vascular structure during interventional procedures. Consequently, the success of these processes heavily depends on the expertise and discernment of practitioners.

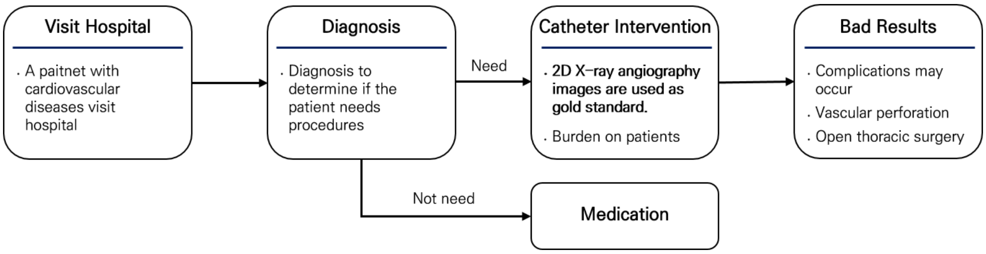

Figure 1. The process for interventional procedures. 2D X-ray angiography images are the primary imaging modality.

Furthermore, during the procedure, acute circumstances, such as an increased heart rate or patient obesity, can obscure the clear identification of vascular structures. Such impediments are notably pronounced during catheterization, hindering the accurate placement of wires through the lesion region. Misdirection of the wire into narrower or frail branches poses a risk to the vessels, precipitating complications. These includes vessel perforation, haemopericardium and cardiac tamponade. In extreme cases, it potentially requires thoracotomy or even leads to patient mortality.

Above all, when facing a chronic total occlusion (CTO) of the coronary artery, the inability of the contrast agent to traverse the occlusion complicates the identification of the distal occluded vessel's pathway. As a result, there is an increased incidence of vascular perforations and emergent scenarios stemming from extensive wiring manipulation.

To mitigate aforementioned ambiguities, researchers have investigated multi-modal registration techniques. These methodologies entail the distinct segmentation of coronary arteries in both 2D and 3D imaging modalities, followed by their subsequent alignment, or the extraction of important features for correspondence matching.

1. A Review on Multi-modal Registrations

A variety of registration-based methodologies have been proposed to address the inherent uncertainties present in 2D X-ray images through feature matching between 2D X-ray angiography images and 3D computed tomography (CT) images [5]–[10].

In [5], statistical motion models of coronary arteries were introduced, based on 4D CT angiography (CTA). Consequently, a 2D/3D + t coronary artery registration method was proposed in [6], utilizing motion models grounded in cardiac and respiratory data. Additionally, Zhu et al. [7] introduced an iterative closest graphs-based matching method utilizing a coarse-to-refine vessel matching mechanism for both rigid and non-rigid transformations.

Recently, convolutional neural network (CNN) models have been utilized in [8]–[10]. They focus on extracting coronary artery centerlines from 2D X-ray images, accompanied by an energy function-based 3D deformation strategy to enable real-time registration.

However, the utility of such approaches is constricted to scenarios where auxiliary 3D CT images can be acquired from patients. Even after the successful extraction of central lines from both modalities, registration remains technically challenging and computationally expensive. This complexity makes these approaches less feasible in exigent clinical emergency situations. Furthermore, these methods are inherently limited in that they are highly dependent on segmentation performance of deep models learned using 2D X-ray and 3D CT images.

2. A Review on Coronary Artery Segmentation in 2D X-ray angiography images

Despite the challenges presented by cardiac motion artifacts and image noise [11], many attempts are undertaken to accurately delineate coronary arteries which are characterized by their elongated and complex tree-like structures in 2D X-ray angiography images.

In early literature, vessel segmentation heavily leveraged Hessian matrix-based vesselness filters, which primarily used the inherent geometrical features of vessels. The Frangi filter [12] is a renowned method that engages a combination of eigenvalues extracted from the decomposition of a Hessian matrix and multi-scale convolution operators.

Subsequently, there have been variants of Hessian-based methods as multi-scale vessel enhancement filters [13]–[17]. Nevertheless, these methods are solely reliant on image information and the susceptibility of second-order derivatives to image noise made the accurate segmentation challenging.

The active contour model exemplifies model-oriented techniques that utilize localized edge information to steer the evolution of curves. It defines the boundaries of vessels by minimizing both internal and external energy functions [18]. It has undergone refined and applied to the segmentation of coronary arteries [19], [20]. In recent studies, the active contour model has been integrated as a loss function within deep neural networks (DNNs) for medical image segmentation, taking into account factors like the segmented contour's boundary length and region fitting [21], [22].

CNNs, a variant of deep neural networks (DNNs), have exhibited its remarkable performance on in the realm of medical image analysis. They automatically extract intricate feature, addressing a main challenge that conventional manual techniques struggle with. U-Net [23], a widely used CNN-based architecture in medical imaging, has spurred the development of various improvements [24]. In contrast to conventional vesselness filtering techniques applied to 2D X-ray images, U-Net presents notable improvements in vessel segmentation tasks. Furthermore, researchers have made many attempts to improve the performance of U-Net in terms of both similarity and accuracy [25], [26].

3. A Review on Inpainting-based Image Restoration

Despite the presence of cardiac motion artifacts and noise in 2D X-ray angiography images, it is highly valuable to acquire accurately segmented coronary arteries, especially in complicated structures. Such precision provides clinicians with an in-depth understanding of the vascular structure during their catheter procedures.

However, from a different perspective, a model capable of reconstructing poorly visible or ambiguous vascular regions affected by cardiovascular disease can also be of clinical importance, as it can provide the operator with the reconstructed image that allows the operator to clearly figure out the vascular region.

A Generative Adversarial Network (GAN) [27] generates realistic data by employing two players—a generator, denoted as G, and a discriminator, denoted as D, in an adversarial manner. It has achieved significant success across a wide range of tasks, including image synthesis [28], image translation [29], image inpainting [30]–[32], and in medical imaging reconstruction [33]–[35]. Especially, image inpainting aims to recover the original image to be realistic from corruption.

Classicial inpainting - Conventional methods restore holes or missing regions in images solely based on surrounding content. These approaches can be categorized in two-folds. One is diffusion-based method [36], [37], and the other is patch-based method [38]–[40]. The diffusion-based approaches fill missing areas using intact neighboring information. On the contrary, patch-based approaches search patches which are similar to the missing patches and use them to fill the missing regions.

While traditional methods can effectively restore damages such as static backgrounds or repetitive patterns, they have been shown inadequate performance on intricate cases due to their limited grasp of high-level image structures. Consequently, the outcomes often exhibit visual disparities with their surroundings that seem implausible and unsatisfactory.

Learning-based Natural Image Inpainting - In recent years, solutions based on deep learning [30], [31], [41]–[44] have emerged to address this long-standing problem in computer vision. These solutions have shown superior performance when compared to classical non-learning-based approaches. They have demonstrated their capability to generate visually more convincing and satisfactory contents, even in complicated scenarios. By harnessing a large volume of extensive datasets through either supervised or unsupervised techniques [27], these methods learn subtle texture and structural patterns. Subsequently, these learned patterns then facilitate the seamless infill of compromised or absent regions, tapping into features sometimes far removed from the target locales.

The initial learning-based solution for this challenge was presented as the context encoder [30]. This approach adopts an encoder-decoder architecture, incorporating both adversarial loss and Euclidean distance-based loss. Building on this, Iizuka et al. [41] enhanced the quality of synthesized images by integrating dilated convolutions into their encoder-decoder design. They also introduced local and global discriminators to ensure semantic consistency between the generated regions and the original areas. Subsequent advancements included contextual attention module [42], which effectively captures extensive contextual information, further elevating improving the overall performance of the restoration process.

Subsequently, [42], [43] and [45] introduced coarse-to-refine frameworks, each employing an auto-encoder network. Yi et al. [46] also employed a two-stage training approach, which is similar to ma et al. [45] Notably, the latter introduces an attention computing module (ACM) and an attention transfer module (ATM) to generate attention scores and aggregate residuals. They enable the generator to produce more visually plausible, high-quality images.

In conjunction with two-stage pipelines, additional edge information has been exploited to generate more realistic images [31]. A two-stream network has also been proposed by guiding and constraining to learn both texture and structure, achieving more detailed image synthesis [47]. In both of the aforementioned methods, the edges and holes are generated by each generator, bypassing the need for auxiliary objective functions.

Learning-based Medical Image Inpainting - Early investigations into medical imaging inpainting [48]–[50] primarily aimed at eliminating artifacts or irregularities for specific purposes, such as image segmentation and registration.

In subsequent research endeavors, Armanious et al. [33], [51] introduced adversarial learning frameworks designed for MR images, capable of generating missing content within both rectangular and arbitrarily shaped regions. Additionally, [35] leveraged edge information and structural components to produce more authentic

CT and MRI images with multi-scale residual blocks and objective functions.

The remaining outline of this dissertation is as follows. Section II.1 presents our novel post-processing techniques for reconnecting fragmented segments within coronary artery segmentation results using local geometric features. Next, Section II.2 introduces another novel approach for reconstructing partially broken vascular structures through the utilization of vesselness-loss-based multi-scale generative adversarial networks. Subsequently, The experimental results and performance of the proposed methods are outlined in Section III. Finally, Section IV provides a comprehensive discussion and analysis of the findings.

II. METHODS

In this paper, we describe methods for restoring missing coronary artery regions. First, from a segmentation perspective, we propose a method to reconnect partially disconnected vessels in the coronary artery segmentation results, in Secion II.1. Then, from the new angle, we propose a method to restore the coronary arteries that are not visible due to lesions within the images, by directly restoring the 2D X-ray angiography images themselves in Section II.2.

1. Reconnection of Fragmented Parts of Coronary Arteries Using Local Geometric Features

Although the performance of deep-learning-based models is improving, the vessels tend to appear partially fragmented in segmentation results of state-of-the-art models due to the limitations of single CNN model, as shown in Fig. 2(a). Also, in the cases of interrupted vessel structures, the segmentation results are partially broken due to the absence of vascular information which is needed for CNN model to infer vessel regions, as shown in Fig. 2(b). Despite the clinical importance of achieving a connected segmentation result for coronary arteries in catheter procedures, the issue of partially broken segmentation results has been rarely studied.

To address this, in our previous study [52], we proposed a fully-automated reconnection method based on local geometric features. In this paper, we assumed that an ideal method for measuring vesselness would result in a probability close to one in the vessel region and a probability close to zero in the background region. However,

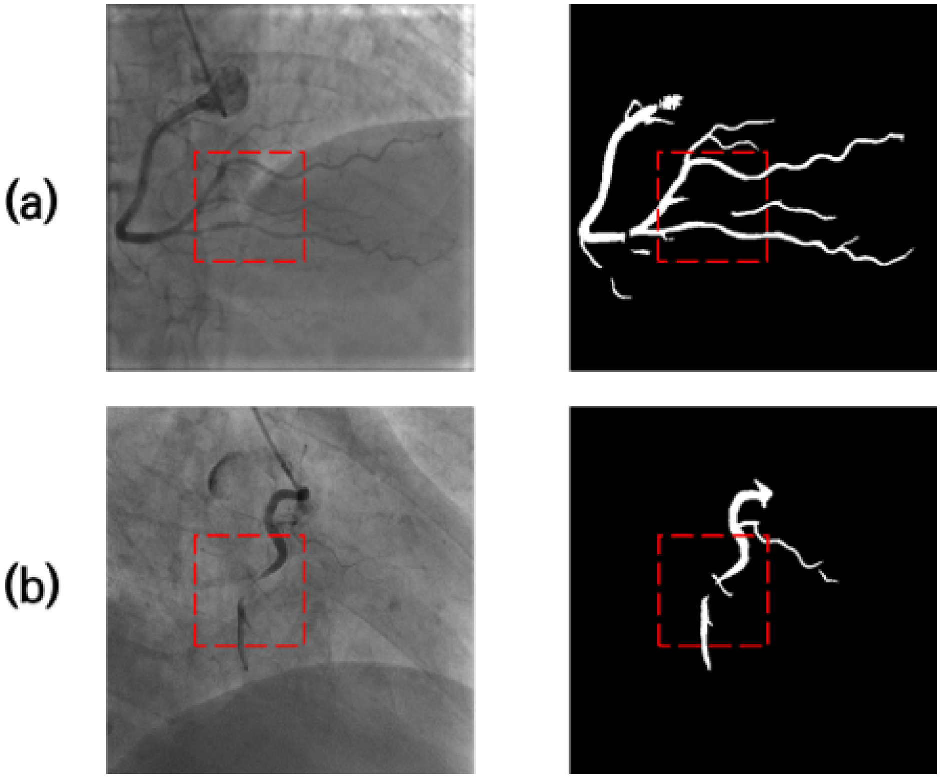

Figure 2. The 2D X-ray images and segmentation results. (a) Partially broken segmentation results occur as a result of limitations in the CNN model. (b) Segmentation results become fragmented when vascular information is absent.

because it is difficult to ensure that consistent probability values are obtained in real-world scenarios, it is important to increase both the accuracy and clinical value by reconnecting the fragmented parts based on the initially predicted masks.

In this paper, we propose a robust method for extracting coronary arteries from X-ray angiographic images by reconnecting the fragmented parts. The complete reconnection process is summarized in Fig. 3. The sophisticated reconnection process is based on a deep neural network-based framework, which results in higher probabilities for only the missing regions. The framework comprises global and local likelihood functions and a geometric prior function.

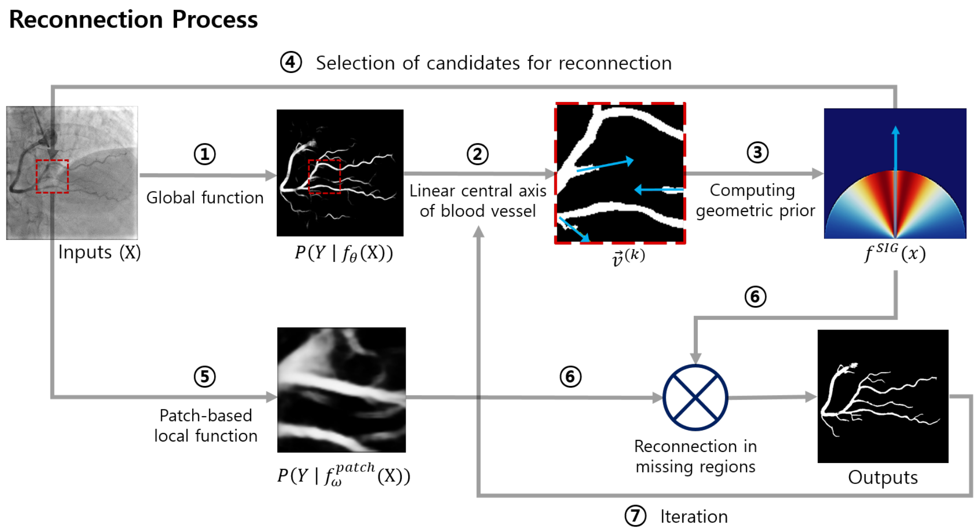

Figure 3. Complete reconnection workflow process for comprehensively identifying the missing regions from the initial segment. First, the initial vessel segment is obtained by global CNN function (step 1), and then all the tip points are detected and the linearly represented central axes of the blood vessels $\vec{v}^{(k)}$ at the tip points are approximated (step 2). To select candidate regions for reconnection, symmetric inverse Gaussian (SIG) function representing the local behavior of the coronary artery is mapped onto the semicircle as geometric priors at the detected tip points (step 3). The regions where the geometric priors overlap each other are selected as the candidate regions (step 4). Local vesselness probabilities are obtained by the auxiliary CNN considering local details at the candidate regions to compensate for the fragments missed in the outputs of the global function (step 5). The local vesselness probabilities are combined with the SIG function to provide the final robust vesselness probabilities (step 6). The reconnection process is iteratively performed until the evaluative values converge (step 7).

Importantly, two expanded version of U-Nets tailored by Res-Net are modeled for local and global functions in our framework. The local function takes the local patch images while the global function takes the entire images as inputs. We trained the auxiliary CNN with the local patch images to provide more robust vesselness probability considering local details, even on the fragments missed in the outputs of the global function. The geometric prior function is separately designed for the forced connection between two disconnected segments using a symmetric inverse Gaussian (SIG) function that reflects the local behavior of the coronary artery.

The proposed method automatically identifies several fragmented locations from the initially predicted masks by the global function, and the patch-based local function and the geometric prior function at the fragmented locations are jointly considered for the posterior vesselness probabilities.

We assume that the entire vessel region C consists of two parts in the X-ray images (Eq. II.1). One is a set of main segments C_{main} from which some parts are missing, and the other is a set of missing parts $C_{missing}$. Therefore, the union of the two parts forms an ideally complete coronary artery as expressed below, where $C_{main} \cap C_{missing} = \phi$.

$$C = \{C_{main} \cup C_{missing}\} \tag{II.1}$$

Many vesselness filtering methods [12], [13] and other deep learning based vesselness methods [23], [25] return the vessel probabilities to segment the entire vessel region C. However, the predicted mask C obtained by vesselness filtering methods often have missing fragments at multiple locations in the coronary artery region. Locating the remaining fragments $C_{missing}$ is a crucial task for clinical applications. Our primary goal is to robustly locate the remaining fragments $C_{missing}$. Overall, our rationale for this task is as follows:

1. Vesselness probability is obtained by a pixel-wise classification using a fully convolutional neural network such as U-Net. However, the outputs have some missing regions even when the CNN model is highly optimized, implying that no single threshold can precisely separate the foreground and background, especially when the target is elongated and thin, such as coronary arteries, as shown in Fig. 4(a).

2. The missing regions need to be reconstructed by local compensation. Considering a local patch whose center point belongs to the coronary artery, the regions of the background and foreground corresponding to the vessel can be nearly balanced. Therefore, the auxiliary CNN learned with the local patch images provides a robust vesselness probability considering local details, even on the fragments missed in previous fully convolutional neural networks.

3. Nevertheless, even the local patch based approach is efficient only when the contrast value of the coronary artery is visible. If noise or other organ parts are present in the local patch, it is difficult to obtain a robust vesselness probability with only patch-based probability by considering local details. In addition, if the contrast value of the coronary artery is not visible, as shown in Fig. 4(b), the patch-based probability goes to zero due to the absence of any local information. Therefore, in fragmented parts of coronary arteries without any contrast information, prior information is assigned based on the surrounding vessel structures using the fact that the curvature of a coronary artery follows a specific distribution.

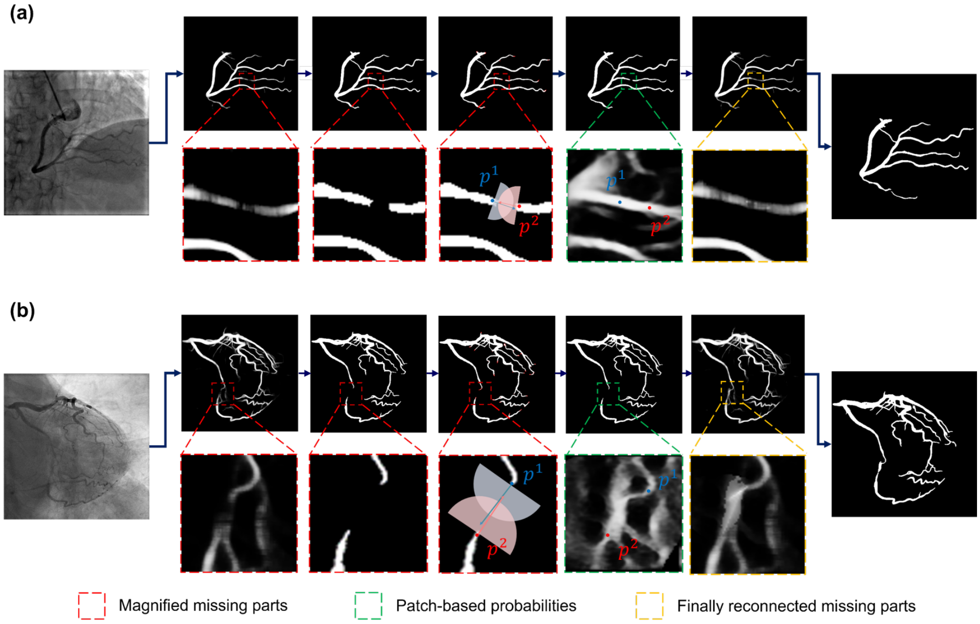

Figure 4. Two cases for which the proposed reconnection framework was applied. The initially segmented coronary arteries with missing parts were finally reconnected by a series of steps. The coronary arteries were not optimally segmented owing to either the limitations in CNN models that could not segment vessels with a single numerical threshold (a), or the lack of information indicating the existence of a vessel (b). To compensate for locations with low probabilities where vessels actually exist, the regions to be reconnected were detected with tip points in vessels and a geometric distribution. Finally, the missing regions were reconnected by compensating the probabilities in the detected regions by considering local patch-based probabilities and distributions.

A. Convolutional Neural Network

CNNs are known to successfully extract image features, and thus can provide vesselness probabilities in X-ray angiography images. U-Net [23] is often used in many studies [53]–[55] to extract coronary arteries from 2D X-ray images and provide a vesselness probability. Furthermore, higher performance can be achieved by customizing the encoding backbone layers [25], [56]–[58]. This leads us to implement Res-Net-50 backbone-encoding based U-Net models ((R)U-Net). It returns the vesselness probability vector by squashing the model output $\mathbf{Y}$ through a softmax function: $P(\mathbf{Y}|\mathbf{f}_\theta(\mathbf{X})) = \text{Softmax}(\mathbf{f}_\theta(\mathbf{X}))$, where $\mathbf{f}$ is the (R)U-Net parameterized by θ. We then classify a pixel $X_{ij} \in C_{\text{main}}$ if $P(Y_{ij} = 1|\mathbf{f}_\theta(X_{ij})) > \tau$ where $\tau = 0.47$ in our implementation. The optimal parameter τ is chosen based on the analysis of the receiver operating curve. We also construct a local patch-based U-Net $\mathbf{f}^{\text{patch}}_\omega(\mathbf{X}_{\text{cropped}})$, parameterized by ω, in order to consider the balance of the background and foreground regions corresponding to vessels. The detail description for the local patch-based U-Net is introduced in Section III.1-1.B-(C).

B. Local Dynamic Prior for Missing Parts

Ideally, C_{main} would have only one entire coronary artery segment, i.e., $C_{\text{missing}} = \phi$. However, C_{main} has multiple sub-segments in most cases, which naturally implies the existence of some missing parts $C_{missing}$.

(A) Linear Approximation of Central Axis of Blood Vessel for Missing Parts

Because the neighboring coronary artery fragment is likely to exist in the central direction of blood vessel, we first determine the linear central axis at the tips of fragmented parts of coronary arteries $\vec{v}^{(k)}$. The steps are as follows. Using a simple CNN, we find the k^{th} tip point $p^{(k)}$ of the fragmented part of the coronary arteries as shown in Fig. 5(a). Because at least two points are essential to define a direction $\vec{v}^{(k)}$, we need another point $q^{(k)}$ that is located slightly further away from the tip $p^{(k)}$ but also inside the vessel region. Then, we define the approximated linear representation of blood vessels as $\vec{v}^{(k)} = p^{(k)} - q^{(k)}$, and its normalized form, $\hat{v}^{(k)}$, where $||\hat{v}^{(k)}|| = 1$. The region growing method is used to identify the point $q^{(k)}$ that is located with the geodesic distance ($\delta = 20$ pixels) from $p^{(k)}$. Note that the distance parameter δ between $p^{(k)}$ and $q^{(k)}$ should be larger than the maximum diameter of coronary arteries. The value (5.0 mm $\approx$ 16 pixels) is referred from the clinical investigation [59]. The examples of tip points and the linear central axes of blood vessels at the tip points are illustrated in Fig. 5(b).

(B) Dynamic Function for Forced Connection

Sometimes, the missing parts are not visible in X-ray images due to severe plaques. In this case, the vesselness probabilities measured by deep neural networks are very low. Therefore, we designed a geometric weight function that enhances the vesselness probability based on the surrounding vascular geometry even when the coronary artery is invisible, thereby representing the natural behavior of the trajectory of blood vessels. Based on centerline ground truth (GT) data, we propose a symmetric inverse Gaussian function

$$f^{\mathrm{SIG}}(\theta; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi(|\theta| + \Delta\theta)^3}} \exp\left[-\frac{\lambda(|\theta| + \Delta\theta - \mu)^2}{2\mu^2(|\theta| + \Delta x)}\right] \tag{II.2}$$

where μ is the mean, and λ and Δx are the shape parameters. In the case of the original inverse Gaussian (IG) function, the probability is zero at $x = 0$, which means that there are no parts with straightness in the local geometry of the coronary arteries. However, we investigated the angle histogram by collecting the angle data between the adjacent directional vectors from the centerlines of coronary arteries, motivated by [60]. This revealed some components that were almost straight.

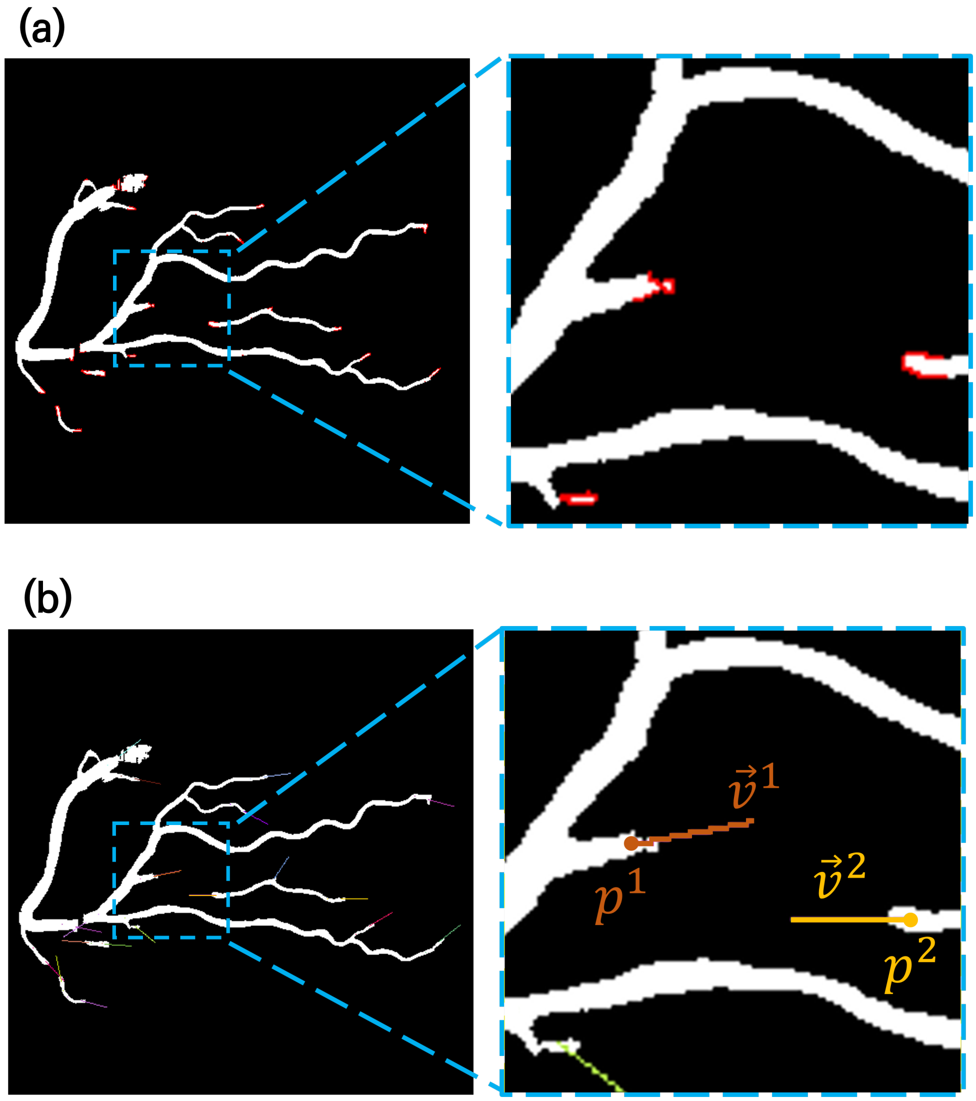

Figure 5. The detected tip points and linearly approximated direction of vessels. (a) Detected multiple tip points (red points) from the binary mask, and magnified views of some disjointed parts. (b) Linearly approximated central axes of blood vessels at the tip points.

For the area in which the weight function is applied, we define the spatial semicircle domain $\Psi^{(k)}$ as

$$\Psi^{(k)} = \left\{(x, y)|x = r\cos\theta, y = r\sin\theta, 0 \le r \le \rho, -\frac{\pi}{2} \le \theta \le \frac{\pi}{2}\right\} \tag{II.3}$$

where ρ varies according to the type of model. The semicircle domain $\Psi^{(k)}$ in Eq. (II.3) is positioned at the missing region $p^{(k)}$ in the direction $\vec{v}^{(k)}$ by applying the linear transformation as follows:

$$\begin{aligned}\Psi'^{(k)} &= T \cdot R \cdot \Psi^{(k)} \\ &= \begin{bmatrix} 1 & 0 & p_x^{(k)} \\ 0 & 1 & p_y^{(k)} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta^* & -\sin\theta^* & 1 \\ \sin\theta^* & \cos\theta^* & 1 \\ 0 & 0 & 1 \end{bmatrix} \Psi^{(k)}\end{aligned} \tag{II.4}$$

where $\theta^* = \cos^{-1}\frac{|\hat{x}\cdot\vec{v}^{(k)}|}{|\hat{x}|\cdot|\vec{v}^{(k)}|}$, and $\hat{x}$ is the unit vector of x. Then, the probabilities from the local-patch U-Net in the domain $\Psi'^{(k)}$ (e.g., $P(Y_{ij}|\mathbf{f}_\omega^{\text{patch}}(X_{ij})), X_{ij} \in \Psi'^{(k)} \setminus C_{\text{main}}$) are multiplied by the symmetric inverse Gaussian (SIG) function in Eq. (II.2) to reflect the local properties of the coronary arteries as shown in Fig. 6.

C. Reconnected Vesselness Map V

We now describe the process for achieving a local patch. From Fig. 4, we can notice that there would be an overlapping area between the two semicircle domains

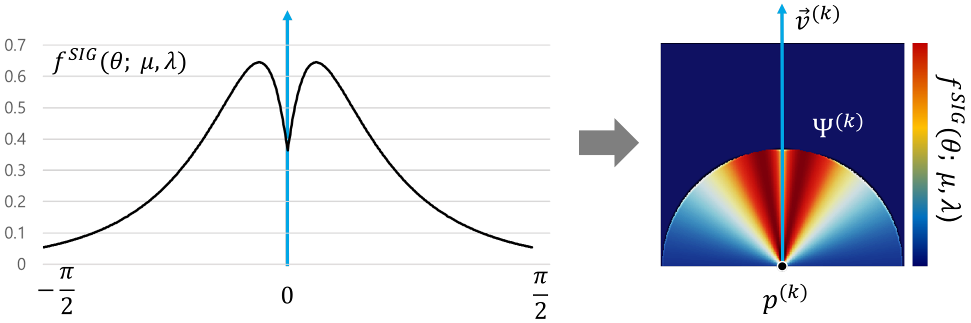

Figure 6. The parameterized symmetric inverse Gaussian. SIG distribution having μ as mean and λ, Δx as shape parameter and reflecting the axis direction of the coronary artery. Here, ρ indicates the radius (half length) of the semicircle, which reflects the SIG distribution.

Ψ^i_j obtained with the linear central directions of the blood vessels at the two tip points. Because the points in this overlapping region are likely to be located in the middle of the disconnected vessels, we created multiple 64×64 patches centered at these points. As these boxed regions may overlap multiple times, we obtain multiple local patch probabilities from the local-patch U-Net. Therefore, we averaged the multiple probabilities output from the local-patch U-Net. Note that this area depends on ρ, the analysis of which is described in the experiment in Section III.1-1.B-(B).

D. Iterative Scheme

The reconnection process is iteratively performed using Algorithm 1. From the initial segment S, all the missing parts in the entire area are gradually reconnected over some iterations. The iteration process is performed until the evaluative values converge, which means that there is no fragmented part in the segmentation result. For example, Fig. 7 shows that the disconnected parts at two locations from the initial result are clearly reconnected over three iterations. In the experimental section, we demonstrate that the algorithm converged sufficiently in approximately three iterations with increase of similarity and accuracy.

Algorithm 1 Reconnection process

V $\leftarrow P(Y_{ij=1}|\mathbf{f}_\theta(X_{ij}))$
Obtain semicircle domain $\Psi'^{(k)}$ from **V**
if $X_{ij} \in \Psi'^{(k)} \setminus C_{\text{main}}, \forall i, j$ **then**
  repeat
    $V_{ij}^{\text{patch}} \leftarrow P(Y_{ij} = 1|\mathbf{f}_\omega^{\text{patch}}(X_{ij}))$
    $V_{ij} \leftarrow V_{ij}^{\text{patch}} f^{\text{SIG}}(X_{ij})$
    Achieve new semicircle domain $\Psi_j'^{(i)}$ from **V**
  until No update on **V**
end if
return V

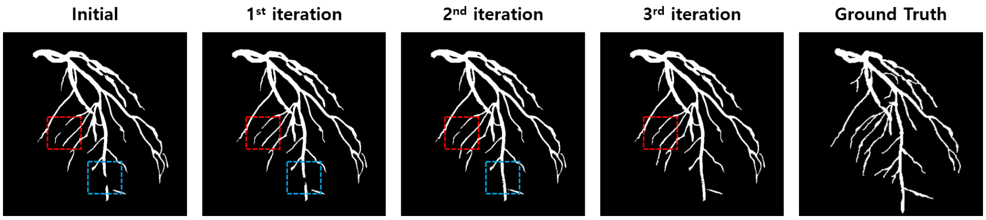

Figure 7. Iterative reconnection process. The completely reconnected final segmentation result after third iteration from initial segment.

2. Reconstruction of Partially Broken Vascular Structures via Vesselness-loss-based Multi-Scale Generative Adversarial Network

The previously introduced post-processing method allowed us to reconnect fragmented coronary artery segmentation results into a single vascular structure. However, this approach not only required time to obtain the initial coronary artery segmentation results from the global network, but also required additional time to obtain patch segmentation results from a local network for localized corrections.

Due to these challenges, we aimed to arternative our previous research from segmentation to image restoration in a new perspective. Particularly, our focus was on reconstructing vascular regions that appear disconnected or blurry due to lesions. In this regard, we introduced a novel approach in [61] that involves restoring the 2D X-ray images themselves.

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 [31], [47] the inpainting model may suffer from being significantly affected by background information, resulting in sub-optimal performance.

Meanwhile, in our previous work [52], 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. II.5:

$$X_i^{\text{blocked}} = X_i^{\text{patch}} \odot (1 - \text{M}) + \text{M} \tag{II.5}$$

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. II.6.

$$\tilde{X}_i^{\text{blocked}} = X_i^{\text{patch}} \odot (1 - \text{M}) + G(X_i^{\text{blocked}}) \odot \text{M} \tag{II.6}$$

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}.

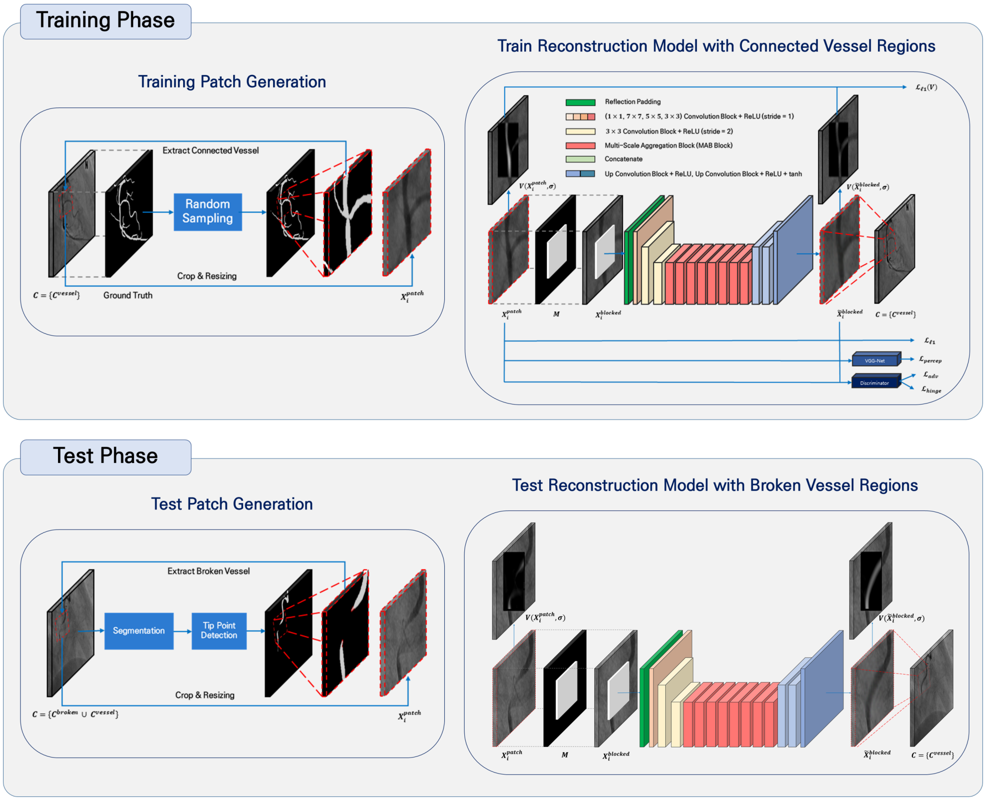

Figure 8. Fully automatic vessel information reconstruction workflow. *Training Phase* : X-ray patches X_i^{patch} containing ideal connected coronary arteries $C = \{C^{\text{vessel}}\}$ are extracted via random sampling of their corresponding ground truths. Next, extracted patches are synthetically 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. *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 [52] perform segmentation masking on the cases, followed by tip point detection to find X_i^{broken} containing broken coronary arteries $C = \{C^{\text{broken}} \cup C^{\text{vessel}}\}$ with a sufficient size to cover broken coronary areas C_i^{broken}. Then, our trained model receives those areas and outputs results $\tilde{X}_i^{\text{blocked}}$ that are similar to the ideal ones C^{vessel}. This phase is rephrased as follows.

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. 8 provides an intuitive illustration of the training and testing pipelines, and Fig. 9 depicts a core module of our MAB network.

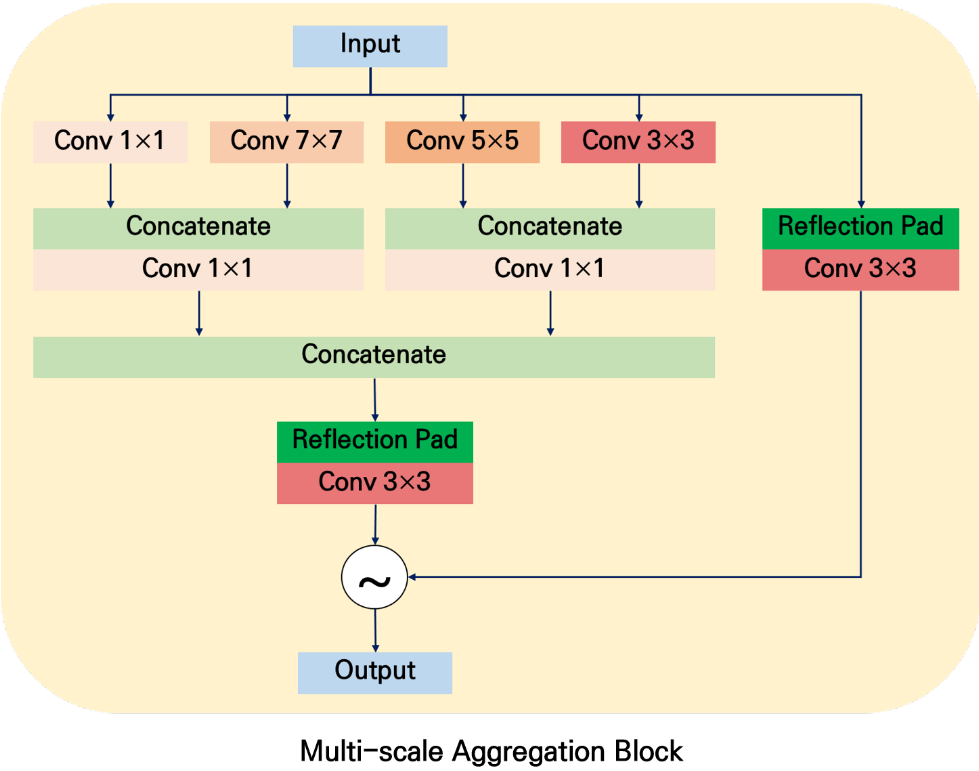

Figure 9. Multi-scale aggregation block (MAB). MAB incorporates convolutional layer with large kernel size into multi-scale fusion blocks.

A. Architecture

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. 8. 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.

(A) Generator

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 [42], [44], 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 [62]–[64] and multi-scale fusion blocks [32], [47] 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. 9.

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 of ResNet [65], we add information flow from another convolutional layer with kernel size 3 to facilitate training and amplify ensemble effect [66] during restoration.

Here, we adopt the gated residual connection strategy, proposed in [32], weighted sums are obtained using gated values in place of simple residual summation.

(B) Discriminator

The discriminator used in this task is obtained from Spectral-Normalized Markovian PatchGAN (SN-PatchGAN) [67], [68], following previous studies [32], [44], [46]. 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.

(A) Hinge Loss

Lim et al. [69] 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. (II.7) and (II.8) describe the generator and discriminator, respectively:

$$\mathcal{L}_G = -\mathbb{E}_{x \sim P_{\tilde{X}^{\text{blocked}}}}[D(\tilde{X}_i^{\text{blocked}})] \tag{II.7}$$

$$\begin{aligned}\mathcal{L}_D = & E_{x \sim P_{X^{\text{patch}}}}[ReLU(1 - D(X_i^{\text{patch}})] \\ & + E_{x \sim P_{\tilde{X}^{\text{blocked}}}}[ReLU(1 + D(\tilde{X}_i^{\text{blocked}})]\end{aligned} \tag{II.8}$$

(B) 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 $\ell 1$ loss function instead of the $\ell 2$ 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. (II.9) as follows:

$$\mathcal{L}_{\ell 1} = ||\tilde{X}_i^{\text{blocked}} - X_i^{\text{patch}}||_1 \tag{II.9}$$

(C) Structural Similarity Index Loss

The structural similarity index measure (SSIM) [70], as shown in Eq. (II.10), is a metric that is used to measure image quality. Unlike $\mathcal{L}_{\ell 1}$ in Eq. (II.9), 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.

$$SSIM(X,Y) = \frac{2\mu_X\mu_Y + C_1}{{\mu_X}^2 + {\mu_Y}^2 + C_1} \cdot \frac{2\sigma_{XY} + C_2}{{\sigma_X}^2 + {\sigma_Y}^2 + C_2} \tag{II.10}$$

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. (II.11).

$$\mathcal{L}_{SSIM} = 1 - SSIM(X,Y) \tag{II.11}$$

(D) Perceptual

Perceptual loss [71] 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 [72] models. For this purpose, we adopt the relu5_1 layer of VGG19.

$$\mathcal{L}_{per} = \frac{1}{C_i H_i W_i} ||\phi_i^{vgg}(\tilde{X}_i^{\text{blocked}}) - \phi_i^{vgg}(X_i^{\text{patch}})||_1 \quad \text{(II.12)}$$

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 $\ell 1$ loss for the reasons described earlier.

(E) Enforced Reconstruction of Vessel Information

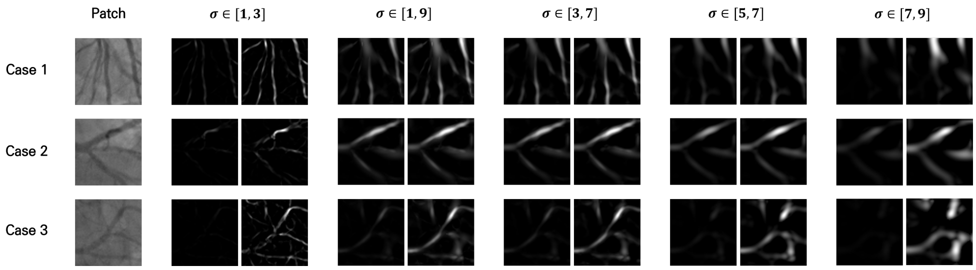

Figure 10. Results of Hessian-based vesselness probability map and normalization depending on the scale, σ, indicating a trade-off relationship. When σ is low, vessel-like structure region has high vesselness probabilities, and thus thin vessels are well segmented; however, vessel-like false positives 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.

In Eq. (II.9), 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. (II.9), 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, $\mathbf{H}$, represents the curvature characteristics of a function by Eq. (II.13), 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, ν, 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.

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 I(x,x)}{\partial x^2} & \frac{\partial^2 I(x,y)}{\partial xy} \\ \frac{\partial^2 I(y,x)}{\partial yx} & \frac{\partial^2 I(y,y)}{\partial y^2} \end{bmatrix} \tag{II.13}$$

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 [12], [73]–[75]. Among them, Frangi et al. [12] 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 (II.14).

$$v(p,\sigma) = \begin{cases} 0 & if \lambda_2 > 0, \\ \exp(-\frac{R_b^2(p,\sigma)}{2\beta^2})(1 - \exp(-\frac{S^2(p,\sigma)}{2C^2})) & \end{cases} \tag{II.14}$$

In equation (II.14), 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, $\mathbf{H}$, varies with respect to the scale, σ, of the matrix, as illustrated in Fig. 10. 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 σ = 3–7 to cover both micro-vessels and thick blood vessels elaborately. Then, using $\tilde{X}^{\text{patch}}$, 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 (II.15).

Then, using $\tilde{X}^{\text{patch}}$, 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. (II.15).

$$\mathcal{L}_{\ell 1(V)} = ||V(\tilde{X}_i^{\text{blocked}}, \sigma) - V(X_i^{\text{patch}}, \sigma)||_1 \tag{II.15}$$

C. Linear Combination of Loss Functions

The five loss functions, introduced in the previous sections, are now summed linearly with weighted constants. Eq. (II.16) and (II.17), describe the objective terms for the discriminator, D, and generator, G, respectively.

$$\mathcal{L}_{D_{total}} = \lambda_1 \mathcal{L}_{hinge} \tag{II.16}$$

$$\mathcal{L}_{G_{total}} = \lambda_1 \mathcal{L}_{hinge} + \lambda_2 \mathcal{L}_{\ell 1} + \lambda_3 \mathcal{L}_{SSIM} + \lambda_4 \mathcal{L}_{per} + \lambda_5 \mathcal{L}_{\ell 1(V)} \tag{II.17}$$

III. EXPERIMENTS AND RESULTS

In this section, we describe the experiments conducted for each method outlined in Sections II.1 and II.2, aimed at restoring partially missing coronary arteries both in segmentation results and 2D X-ray angiography images. The dataset used and the evaluation comparisons are presented in Sections III.1 and III.2, respectively.

1. Experiment for Reconnection of Fragmented Parts of Coronary Arteries Using Local Geometric Features

A. Dataset

The proposed method was trained and evaluated on 3,084 2D X-ray angiography images from 85 patients with cardiovascular diseases. A total of 2,424 images from 65 patients were used as the training dataset, and 660 images from 20 patients were used as the test dataset. There was no significant difference in intensity distribution between the training and test datasets.

Each image was reconstructed to 512×512 pixels, with pixel size ($s_x \times s_y$) ranging between 0.28 mm $\times$ 0.28 mm and 0.36 mm $\times$ 0.36 mm. The number of image frames per patient varies from 12 to 66 serial frames containing moving coronary arteries.

The dataset and corresponding ground truth (GT) masks were provided by Severance Hospital, Yonsei University College of Medicine, South Korea. GT masks were manually annotated by clinical experts with more than five years of experience. The details of the dataset used in the experiment are shown in Table 1.

The experiment was performed on a personal computer with an Intel® Core™ i7-3770K CPU @ 3.50GHz, 32.0 GB RAM, NVIDIA® TITAN X (12.0GB), and was run on Tensorflow 1.9. The refinement process takes 2.8 sec on average per image.

Table 1. Details of experimental dataset for reconnection models

	#Patient	#Frame	# Frame/Patient
Dataset (All)	85	3,084	[19, 66]
Training	65	2,424	[19, 66]
Test	20	660	[23, 53]

Note: The dataset and corresponding ground truth (GT) data are provided from Severance Hospital, Yonsei University College of Medicine, South Korea

B. Quantitative and Qualitative Comparison with Existing Methods

(A) Vascular Tip Detection

The tip points $p^{(k)}$ are described in Section II.1-1.B-(A) and are detected by a simple CNN trained with local binary masks of coronary arteries, as shown in Fig. 11. For labeling task, the patch images are labeled as tip if their centroids are located at the ends of vessels Fig. 11(a) while the patch images are labeled as non-tip if their centroids are located elsewheref Fig. 11(b).

The architecture takes 20×20 input images, two convolution layers with 64 and 128 channels, and three fully connected layers with 1024, 256, 2 neurons for binary classification. Using the trained CNN, we can search for the tip points on the edges of the predicted binary masks. The tip detection task can be solved robustly and is

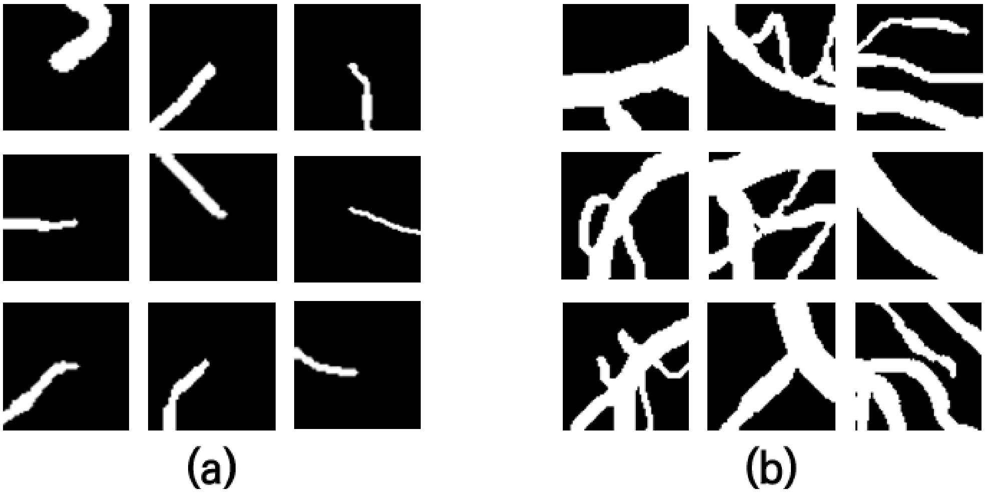

Figure 11. Binary masks for the tip point and non-tip point. From the local binary mask of the initial prediction, tip points of coronary arteries are simply classified using a CNN based tip detector. The local masks cropped at tip area (a) and the local masks cropped at the other area (b) have distinct features.

not sensitive in our method.

(B) Parameters for Dynamic Function

The SIG function described in Section II.1-1.B-(B) considers the curvature property of the coronary arteries. The values of the SIG function obtained by varying θ are mapped onto the corresponding semicircle domains whose centers are at each tip of the fragmented parts. In order to place the semicircles over the missing part where the coronary artery actually exists, the parameters Δ, μ and λ in Eq. (II.2) used to define the SIG distributions were set to 0.4, 1.5, and 3, respectively.

The most sensitive parameter in our system is ρ, used for the local dynamic prior in Eq. (II.3). The reconnection candidate points may not be included when ρ is

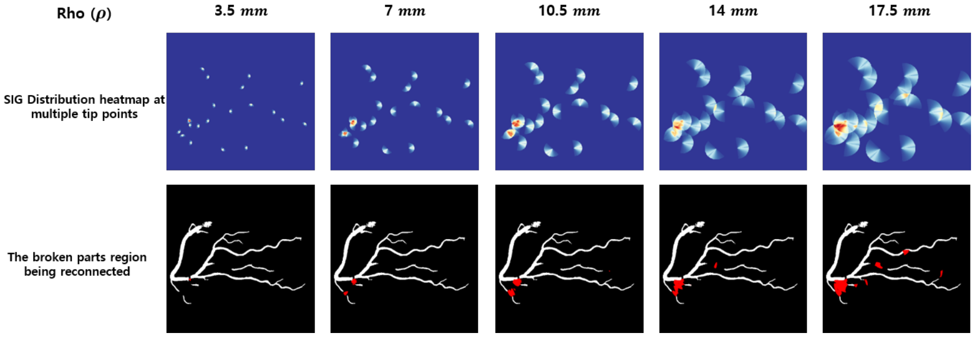

Figure 12. The performance of the reconnection process varied with the semicircle parameters. The semicircle parameters can affect the performance of the reconnection process when the symmetric inverse Gaussian function is applied. The missing regions that need to be reconnected are obtained by the SIG distribution at multiple tip points in the segmentation result. If the radius (ρ) of the SIG distribution sprayed in a semicircle from several tips is too small, there are no overlapping parts. By contrast, if the radius is too large, there are many overlapping parts, including false positive candidates.

too small. On the other hand, some false positive reconnection candidates may be included when ρ is too large. As both these cases must be avoided, it is important to determine the optimal value of ρ.

We analyzed ρ on three models: U-Net, (D)U-Net, and (R)U-Net. The lengths of the missing parts can vary depending on the initial segmentation performance of each model.

Fig. 12 illustrates how the reconnection candidates are chosen by varying ρ based on U-Net. In the case of $\rho = 1.5mm$, the semicircle shapes of local dynamic prior drawn at the two tips of the disconnected fragments are too short to meet each other. This would cause the candidate regions to not be included in the reconnection process. In the case of $\rho = 17.5mm$, on the other hand, the local dynamic prior drawn at the two tips of the disconnected fragments are large enough to meet each other. However, unnecessary candidate regions are also included in the reconnection process. In our experiment, the best segmentation performance for each model is at $\rho = 13mm$ for U-Net, $\rho = 5mm$ for (D)U-Net, and $\rho = 5mm$ for (R)U-Net, respectively.

(C) Local Patch-based U-Net

The architecture for local patch images in Fig. 13 is a compact version of the original U-Net that aims to learn with the small patches. In order to construct the training dataset, 64×64 sized patches are cropped and collected based on the randomly sampled center locations on the vessel region from labeled masks. Importantly, the collected patches have the balance of the background and foreground regions corresponding to vessels, thus the model can predict the vesselness probabilities considering local details. The patch-based U-Net are utilized to reconnect the disconnected fragments from the initially predicted segment.

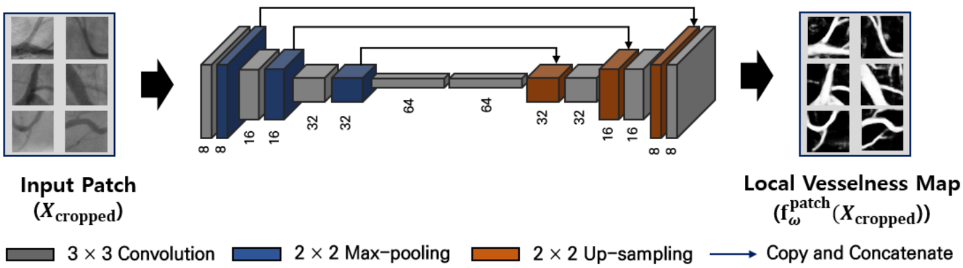

Figure 13. The local patch-based CNN network. The network is a compact version of the original U-Net that befits learning with the small patches (X_{cropped}), and the local vesselness map is obtained by the trained model ($f_{\omega}^{\text{patch}}(X_{\text{cropped}})$).

(D) Evaluation Metrics

Using similarity metrics, the proposed methods can be evaluated and compared with other methods. The Dice similarity coefficient (DSC) and Jaccard Index (JI) are the most referenced metrics for the measurement of similarity in the literature. Both metrics are calculated pixel-wise by comparing the outputs of each method with the ground truth. Detailed descriptions are shown in Table 2.

Table 2. Metrics and descriptions for evaluating proposed methods

Metrics	Description
Dice similarity coefficient (DSC)	2TP / (2TP + FP + FN)
Jarracd index (JI)	TP / (TP + FP + FN)

Note: Let X and Y denote the outputs of an algorithm and the ground truth, respectively; TP:=$|x \in Y|$; FP:=$|x \notin Y|$; TN:=$|(X \cup Y)^c|$; FN:=$|Y| - |x \in Y|$

The values of true positive (TP), false positive (FP), true negative (TN), and false negative (FN) are determined by whether an image pixel corresponds to a vessel in both ground truth and segmented results, only in one of them, or corresponds to the background in one of them but a vessel in the other. TP, FP, TN, and FN are defined as the number of vessel pixels classified as vessel, the number of background pixels classified as vessel, the number of background pixels classified as background, and the number of vessel pixels classified as background, respectively.

(E) Evaluation at Entire Image

Before deep learning was actively applied, the Frangi filter and similar image-gradient-based methods were widely used[12], [13], [76] to detect vessel-like structures in images. In addition, U-Net provided a breakthrough in the quality of results achieved and has become a popular neural network method for segmentation purposes in medical images. Further, extended versions of U-Net, whose backbone layers are customized with Dense-Net or Res-Net, exhibit higher performance in medical image analysis. These methods were therefore utilized as baseline models for our quantitative evaluations.

We illustrate the performance of our method with a quantitative comparison in Table 4, based on the evaluation metrics DSC and JI defined in Table 2. Note that the results of three classical methods, Frangi[12], Krissian [13], and IVGMM [76] are referenced from other studies based on different datasets [76]. Thus, we cannot directly compare the proposed methods. Among the image gradient-based methods, IVGMM exhibited the highest DSC score of 0.704.

Table 3. Quantitative comparison with the existing methods

	Solvers	DSC	JI
Image gradient-based	Frangi [12]	0.521	0.358
	Krissian [13]	0.520	0.357
	IVGMM [76]	0.704	0.553
CNN based	U-Net [23]	0.790	0.659
	(D)U-Net [25]	0.810	0.684
	(R)U-Net [25]	**0.811**	**0.685**
Proposed	U-Net + Proposed	0.794	0.664
	(D)U-Net + Proposed	0.811	0.685
	(R)U-Net + Proposed	**0.812**	**0.686**

Notes: The results by Frangi, Krissian, and IVGMM (non-deep learning methods) are referenced from other studies based on a different dataset [76]. Hence, Frangi, Krissian, and IVGMM cannot be directly compared with the proposed methods; U-Net; (D)U-Net, Dense-Net backbone-encoding based U-Net; (R)U-Net, Res-Net backbone-encoding based U-Net; DSC, dice similarity coefficient; JI, Jaccard index; CNN, convolutional neural network

However, by employing a deep learning method (U-Net [23]), the DSC score begins from 0.790, which is approximately 9 percent higher than that of the classical methods. (D)U-Net and (R)U-Net [25] exhibit DSC scores of 0.810 and 0.811, respectively. However, we can observe that all the results from each of the models have some fragmented regions, as shown in Fig. 14(c) and the blue dotted box in 14(d).

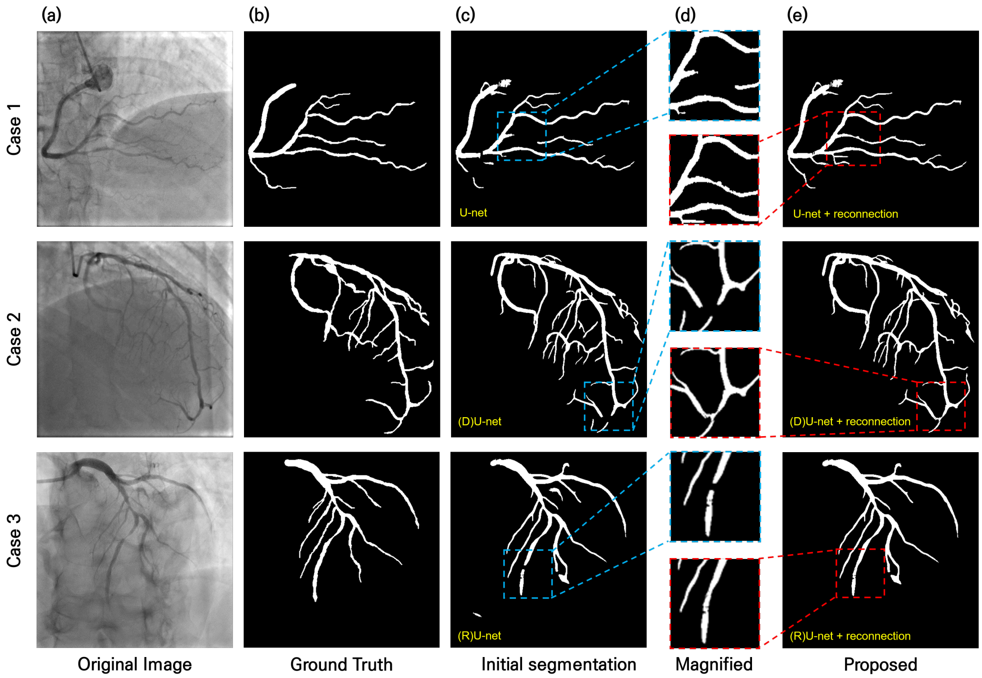

Figure 14. Qualitative results obtained from applying the proposed method to the initial segmentation results obtained by applying three different U-Net models. (a) Input 2D X-ray angiography, (b) Ground truth, (c) Initial segmented coronary arteries containing missing parts from different single models, (d) Magnified missing areas (blue) and reconnected areas (red), (e) Final segmentation results containing complete coronary artery by using the proposed method.

The proposed method was applied to each U-Net-based model to observe the changes in the DSC and JI scores, and how the missing regions are actually reconnected. The results demonstrate improvements for all the three base models, U-Net,

Table 4. The detail number of missing areas on average per patient before and after the refinement

Patient #	# of frames	# of missing areas on average	
		Initial	Refined
1	42	2.810	1.190
2	23	1.174	0.348
3	45	1.600	0.644
4	20	1.350	0.650
5	33	1.970	1.303
6	32	1.656	0.781
7	50	1.200	0.720
8	27	1.481	1.000
9	33	0.848	0.455
10	28	3.393	2.500
11	27	2.556	1.222
12	35	3.143	1.257
13	32	2.719	1.344
14	35	2.171	1.314
15	34	2.765	1.412
16	35	1.714	0.714
17	27	3.630	1.741
18	34	3.382	2.059
19	34	3.147	1.912
20	34	3.441	1.382
Average	33	2.308	1.197

Notes: the initial results are from U-Net model. The range of missing areas per frame is zero to eleven. The number of missing areas on average is reduced by almost half after the refinement.

(D)U-Net, and (R)U-Net in both DSC and JI scores. The best accuracies were 0.812 DSC and 0.686 JI scores from (R)U-Net.

Even though the clinical impact of the reconnection is substantial, there was only a 0.001 difference from 0.811 to 0.812 in the DSC score, because coronary arteries are thin and sparse in all images. This makes it difficult to evaluate the effectiveness of the reconnection over the entire image scope. Therefore, the detail counts of missing areas before and after the reconnection process are presented in Table 4. We found that the number of missing areas on average is reduced from 2.308 to 1.197 after the refinement, which has been reduced by almost half.

In addition, in the next section, we describe an additional experiment to examine the effectiveness of reconnection only in the local region of interest.

(F) Evaluation at Local missing Regions

We also carried out an additional experiment to evaluate the effectiveness of the reconnection method at only the missing regions. Since the missing fragments may exist with different sizes at multiple locations, the measuring areas where the reconnection process was performed are defined using tip points of the disconnected fragments. To define the local region of interest (ROI) at missing regions, dynamic boxed areas are defined by the box parameters x, y, w, h as based on the fact that the missing region exists between two adjacent fragments as,

$$
\begin{aligned}
x &= \min(p_x^{(k')}, p_x^{(k'')}) \\
y &= \min(p_y^{(k')}, p_y^{(k'')}) \\
w &= |p_x^{(k')} - p_x^{(k'')}| \\
h &= |p_y^{(k')} - p_y^{(k'')}|
\end{aligned}
\tag{III.1}
$$

where k' and k'' are two indices for two adjacent tip points from the adjacent fragments. Now, all the local ROIs at missing regions are represented by multiple boxes using the parameters in Eq. (III.1), and then, the DSC and JI scores were evaluated in the defined local ROIs.

For U-Net, (D)U-Net, and (R)U-Net models, DSC scores of 0.280, 0.233, and 0.227 from the initial results increased to 0.600, 0.542, and 0.589, respectively, after the reconnection process. This implies an significant increase of 0.330 in the mean DSC score as shown in Table 5.

Table 5. Evaluation in locally reconnected regions

	Initial		Reconnected	
Metrics	DSC	JI	DSC	JI
U-Net	0.280	0.163	**0.600**	**0.429**
(D)U-Net	0.233	0.132	**0.542**	**0.371**
(R)U-Net	0.227	0.128	**0.589**	**0.417**

Notes: The quantitative results of the single model and proposed methods, respectively, in reconnected regions. The results show that there was significant improvement, evidenced by a more than twofold increase of scores, at local patches around the area where the reconnection took place.

(G) Analysis for Iterative Scheme

The proposed method gradually connects the disconnected parts through the iterative scheme introduced in Section II.1-1.C. The reconnection process via the iterative scheme is shown in Fig. 7. Two disconnected regions (red and blue boxes) were observed in the initial segmentation result. The critical missing regions are completely recovered in the third iteration, thereby achieving convergence. In this case, the similarity measure values did not change significantly even with further iterations.

We also observed how the performance improves quantitatively based on the two similarity metrics of dice similarity coefficient (DSC) and Jaccard index (JI), on three models and over three iterations, as shown in Table 6. Residual U-Net showed the highest performance for all measurements in the third iteration. Compared to the performance of the initial segmentation (DSC=0.811, JI=0.685), the performance after the third iteration (DSC=0.812, JI=0.686) did not seem to significantly increase; however, the difference in clinical value increased significantly owing to the fact that the disconnected region was connected. In our experiment, three iterations were sufficient to reconnect disconnected fragments at the initial mask.

Table 6. Evaluation according to the number of iterations

Iteration	Initial		First		Second		Third	
Metrics	DSC	JI	DSC	JI	DSC	JI	DSC	JI
U-Net	0.790	0.659	0.792	0.662	0.794	0.663	0.794	0.664
(D)U-Net	0.810	0.684	0.810	0.685	0.811	0.685	0.811	0.685
(R)U-Net	0.811	0.685	0.811	0.686	0.811	0.686	**0.812**	**0.686**

Notes: U-Net; (D)U-Net, Dense-Net backbone-encoding based U-Net; (R)U-Net, Res-Net backbone-encoding based U-Net; DSC, dice similarity coefficient; JI, Jaccard index;

2. Experiment for Reconstruction of Partially Broken Vascular Structure

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. 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 $\times$ 512 image with a size between 0.28 mm $\times$ 0.28 mm and 0.36 mm $\times$ 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 $\times$ 128 patches centered on the pixel. When a 64 $\times$ 64-sized blocked mask is positioned at the center of a patch where actual blood vessels are restored, only those 128 $\times$ 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 7. 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.

Table 7. Details of experimental dataset for reconstruction models

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.

B. Quantitative and Qualitative Comparison with Existing Methods

(A) 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 (III.2).

$$PSNR = 10 \cdot \log_{10} \frac{MAX^2}{MSE} \tag{III.2}$$

- **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 (III.3).

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{III.3}$$

- **SSIM** : SSIM can be used not only as an additional loss function to generate an image, as described in Section (A), 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 (II.10).

(B) 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 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. III.4.

$$X_i^{\text{damaged}} = (X_i^{\text{patch}} \odot (1 - \mathrm{M}) + ((X_i^{\text{patch}} \odot \mathrm{M}) \otimes B^n + N^n) \tag{III.4}$$

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. III.5.

$$\tilde{X}_i^{\text{damaged}} = X_i^{\text{patch}} \odot (1 - \mathrm{M}) + G(X_i^{\text{damaged}}) \odot \mathrm{M} \tag{III.5}$$

For fair comparison between the proposed MAB network and other state-of-the-art networks [30], [32], [33], [46], 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: $\lambda_1 = 1$, λ_2 = 256, $\lambda_3 = 1$, $\lambda_4 = 128$ and $\lambda_5 = 256$.

As shown by the quantitative comparison outlined in Table 8, 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

Table 8. 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 [30]	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 + $\mathcal{L}_{\ell 1(V)}$	33.609	0.195	0.902	**32.785**	**0.246**	**0.891**	**31.649**	**0.317**	**0.881**
CE + $\mathcal{L}_{\ell 1(V)} + L_{SSIM}$	**34.051**	**0.178**	**0.907**	32.504	0.257	0.888	31.405	0.330	0.879
IPA [33]	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 + $\mathcal{L}_{\ell 1(V)}$	35.291	0.137	0.922	**33.847**	**0.199**	**0.908**	**33.122**	**0.238**	**0.901**
IPA + $\mathcal{L}_{\ell 1(V)} + L_{SSIM}$	**35.477**	**0.126**	**0.925**	33.711	0.200	0.907	32.895	0.249	0.899
HR [46]	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 + $\mathcal{L}_{\ell 1(V)}$	**34.338**	**0.168**	**0.904**	32.786	0.248	0.887	**32.508**	0.299	0.881
HR + $\mathcal{L}_{\ell 1(V)} + L_{SSIM}$	34.274	0.173	0.903	**32.940**	**0.241**	**0.891**	32.108	**0.294**	**0.883**
AOT [32]	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 + $\mathcal{L}_{\ell 1(V)}$	**35.553**	**0.126**	0.918	33.804	0.194	0.902	32.813	0.253	0.892
AOT + $\mathcal{L}_{\ell 1(V)} + L_{SSIM}$	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 + $\mathcal{L}_{\ell 1(V)}$	35.669	0.120	0.918	33.990	0.186	0.904	33.127	0.231	0.897
MAB + $\mathcal{L}_{\ell 1(V)} + L_{SSIM}$	**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 [30], IPA [33], HR [46], and AOT [32]). 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.

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 [77].

Next, we evaluate the performance improvement of the vessel-specific restoration process achieved using our proposed vesselness-loss function, $\mathcal{L}_{\ell 1(V)}$. Additional experiments are conducted with the regularization of the baseline networks and the proposed MAB network using $\mathcal{L}_{\ell 1(V)}$ or L_{SSIM} or both. As shown in Table 8,

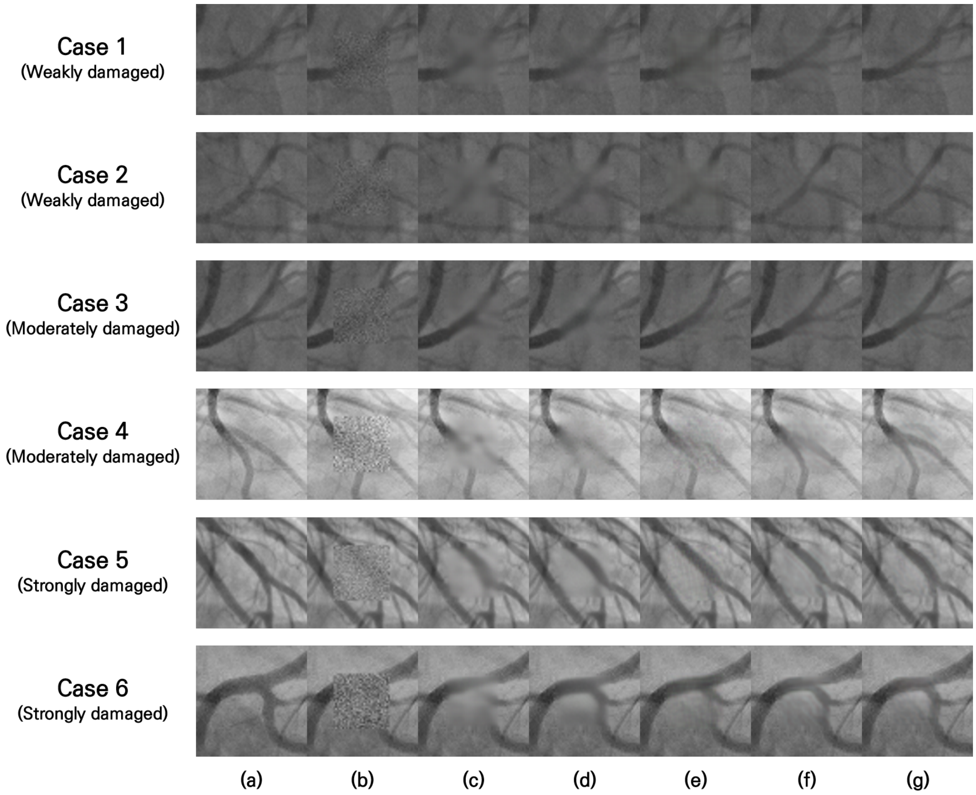

Figure 15. Qualitative results reconstructed using synthesized damaged images. Qualitative results $\tilde{X}_i^{\text{damaged}}$ reconstructed using each baseline network using synthesized damaged images $X_i^{\text{patch-damaged}}$ under different settings: (a) intact X-ray images X_i^{patch} (GT); (b) synthesized images $X_i^{\text{patch-damaged}}$ (Input); (c) CE [30]; (d) IPA [33]; (e) HR [46]; (f) AOT [32]; (g) our proposed MAB network.

adding L_{SSIM} does not always lead to improved vessel-specific reconstruction performance. On the contrary, optimization using $\mathcal{L}_{\ell 1(V)}$ contributes to improvement in vessel reconstruction—every baseline model exhibits its highest quantitative results in all arbitrarily damaged settings. Further, training networks using $\mathcal{L}_{\ell 1(V)}$ 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 $\mathcal{L}_{\ell 1(V)}$ or both $\mathcal{L}_{\ell 1(V)}$ and L_{SSIM} are presented in Fig. 15. The proposed MAB network achieves the highest performance in vessel reconstruction, as demonstrated by the data in Fig. 15 and Table 8.

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 [30], [33], [46] 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 [32] still generates some 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.

(C) 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 9, 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) [32]. 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. 16. From the reconstructed results of [30], [32], [33], [46], 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.

(D) Evaluation of the Proposed Fully-Automatic Reconstruction Methods

The proposed fully automated reconstruction process is evaluated in interrupted vessel structure cases involving broken coronary arteries, $C = \{C^{\text{broken}} \cup C^{\text{vessel}}\}$. The process comprises three steps. First, we utilize our previous studies on ROI detection methods. Thus, tip detection sub-methods are used to search for broken

Table 9. Quantitative comparison of various methods on the reconstruction of completely blocked images

Methods	PSNR ↑	MSE ↓	SSIM ↑
CE [30]	28.809	0.630	0.855
CE + $\mathcal{L}_{SSIM}$	28.995 (+0.186)	0.610 (-0.020)	0.859 (+0.004)
CE + $\mathcal{L}_{\ell 1(V)}$	29.120 (+0.311)	0.603 (-0.027)	0.858 (+0.003)
CE + $\mathcal{L}_{\ell 1(V)} + L_{SSIM}$	**29.148 (+0.339)**	**0.596 (-0.034)**	**0.862 (+0.007)**
IPA [33]	29.190	0.580	0.871
IPA + $\mathcal{L}_{SSIM}$	29.183 (-0.007)	0.590 (+0.010)	0.872 (+0.001)
IPA + $\mathcal{L}_{\ell 1(V)}$	29.277 (+0.087)	0.570 (-0.010)	0.870 (-0.001)
IPA + $\mathcal{L}_{\ell 1(V)} + L_{SSIM}$	**29.365 (+0.175)**	**0.570 (-0.010)**	**0.872 (+0.001)**
HR [46]	29.251	0.600	0.864
HR + $\mathcal{L}_{SSIM}$	29.242 (-0.009)	0.592 (-0.008)	0.863 (-0.001)
HR + $\mathcal{L}_{\ell 1(V)}$	29.421 (+0.170)	0.572 (-0.028)	0.864 (+0.000)
HR + $\mathcal{L}_{\ell 1(V)} + L_{SSIM}$	**29.457 (+0.205)**	**0.564 (-0.036)**	**0.866 (+0.002)**
AOT [32]	29.586	0.557	0.875
AOT + $\mathcal{L}_{SSIM}$	29.700 (+0.114)	0.526 (-0.031)	0.873 (-0.002)
AOT + $\mathcal{L}_{\ell 1(V)}$	29.614 (+0.028)	0.546 (-0.011)	0.870 (-0.005)
AOT + $\mathcal{L}_{\ell 1(V)} + L_{SSIM}$	**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 + $\mathcal{L}_{\ell 1(V)}$	**29.850 (+0.190)**	**0.515 (-0.019)**	**0.873 (+0.000)**
MAB + $\mathcal{L}_{\ell 1(V)} + L_{SSIM}$	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[30], IPA[33], HR[46] and AOT[32]). Furthermore, optimizing the vesselness-loss function leads to performance improvements in all networks.

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, $C = \{C^{\text{broken}} \cup C^{\text{vessel}}\}$, relative to the

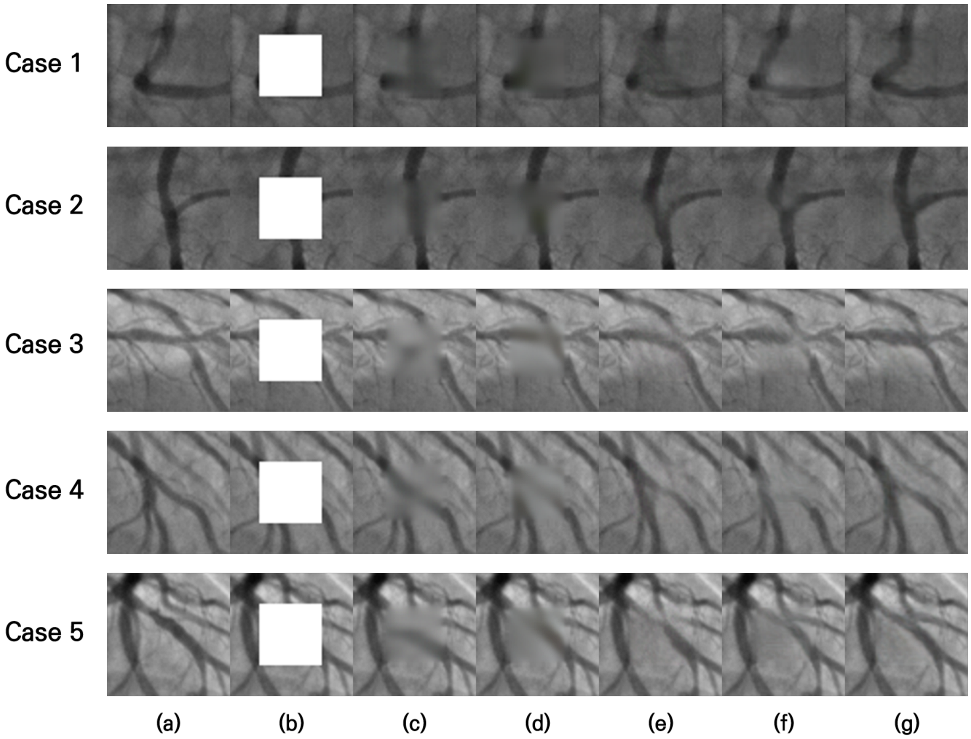

Figure 16. Qualitative results reconstructed using blocked images. Qualitative results $\tilde{X}_i^{\text{blocked}}$ of baseline networks using completely blocked images X_i^{blocked}: (a) ideal patches X_i^{patch} containing intact vessel information; (b) blocked patches X_i^{blocked} without vessel information; (c) CE [30]; (d) IPA [33]; (e) HR [46]; (f) AOT [32]; (g) proposed MAB network.

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. 17. 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

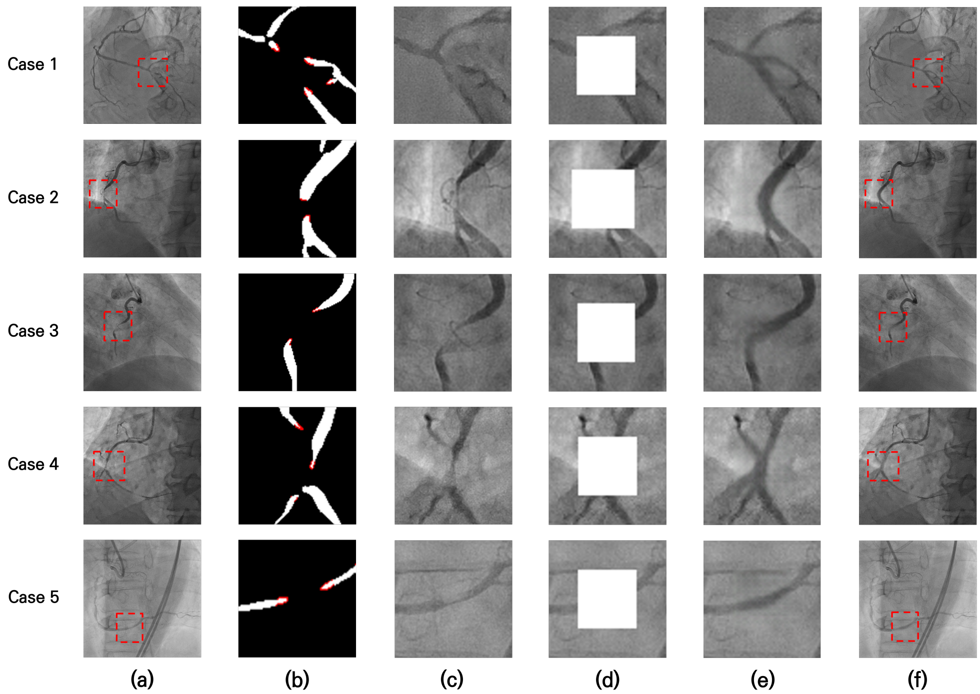

Figure 17. Results of the proposed fully automated reconstruction process in X-ray images. (a) 2D X-ray images that contain broken coronary arteries $C = \{C^{\text{broken}} \cup C^{\text{vessel}}\}$; (b) automatically detected ROI areas via ROI detection methods, with tip points colored in red; (c) 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 $C = \{C^{\text{vessel}}\}$.

the explicit guidance provided by our vesselness-loss.

(E) 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

8 and 9 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 $\mathcal{L}_{\ell 1(V)}$ and L_{SSIM}, 2) with only L_{SSIM}, 3) with only $\mathcal{L}_{\ell 1(V)}$, and 4) with both $\mathcal{L}_{\ell 1(V)}$ and L_{SSIM}.

Figs. 18(a) and (b) present an X-ray image and the synthesized blocked images. The model trained without $\mathcal{L}_{\ell 1(V)}$ and L_{SSIM} generates blurry results, as depicted in Fig. 18(c). When trained using only L_{SSIM}, the model yields visually improved results, as depicted in Fig. 18(d), compared to those depicted in (c). On the other hand, for the models trained using $\mathcal{L}_{\ell 1(V)}$, the edges of the vessel regions are clearly reconstructed, as shown in Fig. 18(e), compared to those depicted in (c) and even (d). Moreover, the intensity 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

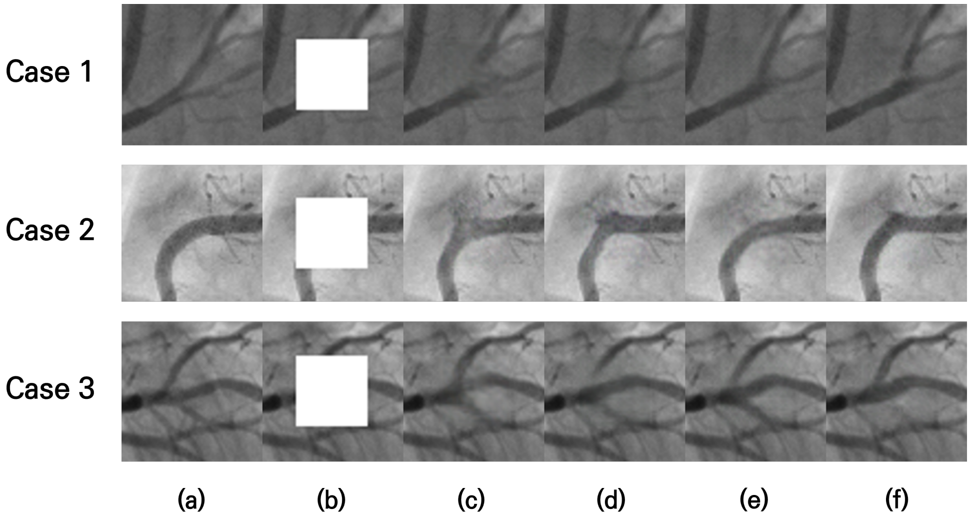

Figure 18. Multiple results from ablation studies to visualize effectiveness of our novel vesselness-loss function on our novel MAB network. (a) X-ray images X_i^{patch}; (b) blocked images X_i^{blocked}; (c) blurry results of model trained without both f the model trained without both ; (d) blurry results 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 $\mathcal{L}_{\ell 1(V)}$; (f) results with sharply represented edges in vessel regions of model trained with both $\mathcal{L}_{\ell 1(V)}$ and L_{SSIM}.

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 $\mathcal{L}_{\ell 1(V)}$ and L_{SSIM} are highly synergetic in vessel reconstruction, as revealed by the comparison between Fig. 18(f) and (d, e)—the former depicts the results obtained using the model trained using both $\mathcal{L}_{\ell 1(V)}$ and L_{SSIM}, whereas the latter depicts those obtained using the model trained using either $\mathcal{L}_{\ell 1(V)}$ or L_{SSIM}.

3. Experiment for Demonstrating Clinical Feasibility

In addition, to demonstrate the clinical feasibility of deep learning based restoration methods, using inpainting GAN, we conducted further experiments for a paper currently under review titled 'Deep Learning Approaches for Image Restoration in Invasive Coronary Angiography.'

A. Dataset

The dataset comprised 2,147 2D X-ray ICA images from 60 patients. To prevent GAN from losing focus during training owing to the sparsely populated coronary arteries in the original 512 × 512 images, we extracted 128 × 128 patches from the major coronary vessels. This allowed us to generate a significantly greater number of patches containing a greater percentage of the coronary artery region within the image. The details of the datasets used for training and testing are presented in Table 10.

Table 10. Overview of the dataset for demonstrating clinical feasibility

Size	Dataset	#Patient	#Frame	# Min Frame / Patient	# Max Frame / Patient
512 × 512	Total	60	2,147	19	66
	Train	35	1,158	19	66
	Test	25	989	22	63
128 × 128	Total	60	22,659	1	39
	Train	35	12,758	1	34
	Test	25	9,901	1	39

Notes: The dataset was obtained from Severance Hospital, Yonsei University College of Medicine, South Korea.

As in Section 2, we created partially degraded patch images denoted as $I_{degraded}$, applying Gaussian blur function $B()$ and noise function $N()$ to real image I_{real}, to access the potential of inpainting GAN network for realistic restoration. However,

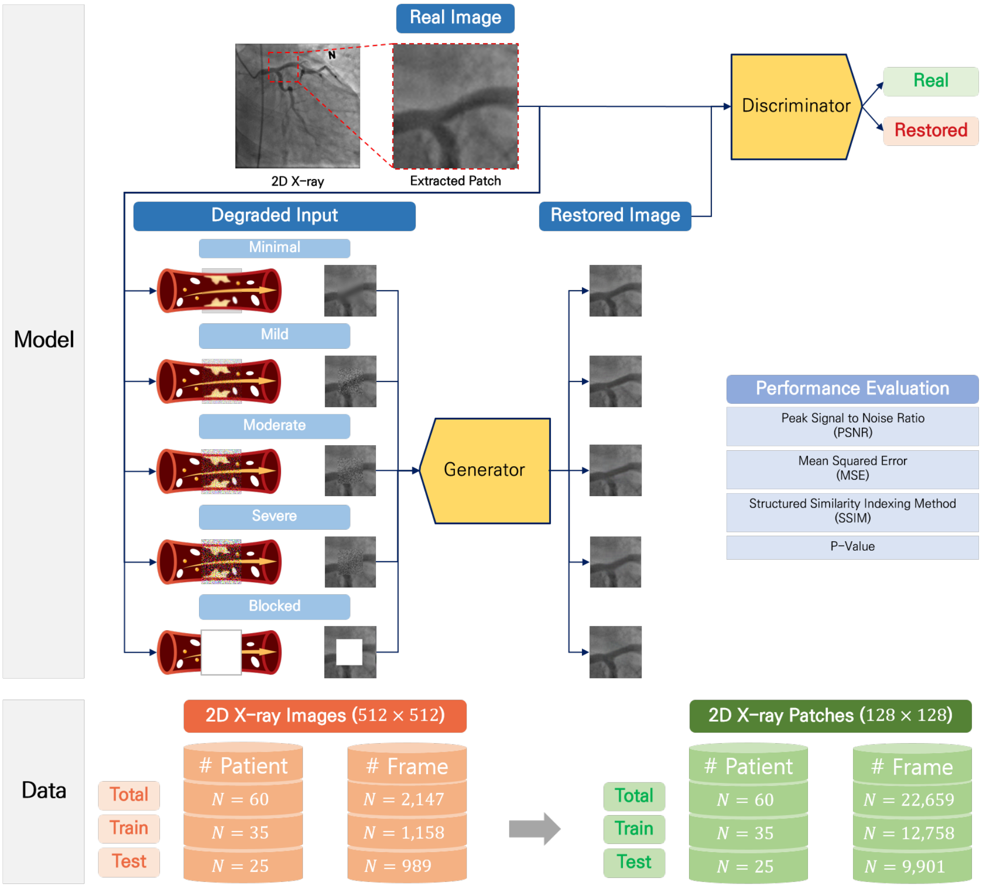

Figure 19. The description of the graphical abstract. The graphical abstract provides an overview the main concept, evaluation metrics, and datasets utilized in this study. The restored results illustrate the practicality of deep learning-based restoration methods in revitalizing a spectrum of scenarios encountered in real clinical environments.

different from Section 2, here we defined the following five severity levels: minimal, mild, moderate, severe, and blocked. The degradation of each level is as follows: Minimal($B^1(I_{real})$), Mild($B^1(I_{real})$, $N^1(I_{real})$), Moderate($B^3(I_{real})$, $N^1(I_{real})$), and Severe($B^5(I_{real})$, $N^1(I_{real})$). Blocked images devoid of vascular information were generated using a function: $Blocked()$. This function mapped the pixel values in a central 64 $\times$ 64 region of the original patch images (I_{real}) to a random pixel value (e.g., 255) while preserving the pixel values in the surrounding regions, as demonstrated in $Blocked(I_{real})$. The graphical abstract of the experiment is shown in Fig. 19 .

Table 11. Quantitative Analysis of Image Quality through PSNR, MSE, and SSIM: A Pre- and Post-Restoration Comparison

	PSNR (n=9,901)			MSE (n=9,901)			SSIM (n=9,901)		
	Before Restoration	After Restoration	P-value	Before Restoration	After Restoration	P-value	Before Restoration	After Restoration	P-value
Minimal($B^1(I_{real})$)	33.27 ± 2.71	37.51 ± 2.71	<0.0001	0.00171 ± 0.00113	0.00060 ± 0.00036	<0.0001	0.868 ± 0.021	0.914 ± 0.019	<0.0001
Mild($B^1(I_{real})$, $N^1(I_{real})$)	27.94 ± 2.87	33.34 ± 2.33	<0.0001	0.00590 ± 0.00374	0.00161 ± 0.00098	<0.0001	0.745 ± 0.013	0.861 ± 0.023	<0.0001
Moderate($B^3(I_{real})$, $N^1(I_{real})$)	27.63 ± 2.90	31.97 ± 2.64	<0.0001	0.00638 ± 0.00413	0.00229 ± 0.00153	<0.0001	0.737 ± 0.012	0.845 ± 0.022	<0.0001
Severe($B^5(I_{real})$, $N^1(I_{real})$)	27.39 ± 2.94	31.46 ± 2.71	<0.0001	0.00678 ± 0.00444	0.00261 ± 0.00183	<0.0001	0.735 ± 0.012	0.844 ± 0.021	<0.0001
$Blocked(I_{real})$	10.99 ± 1.30	28.79 ± 3.03	<0.0001	0.2486 ± 0.06543	0.00504 ± 0.00382	<0.0001	0.718 ± 0.015	0.821 ± 0.023	<0.0001

Notes: PSNR, peak signal-to-noise ratio; SSIM, structured similarity indexing methods; MSE, mean squared error.

As shown in Table 11, in the assessment of PSNR, the scores for the minimally damaged, mildly, moderately, and severely damaged images were comparable, except for those of the blocked images. Upon applying the restoration method, there was a notable increase of up to 12.7%, 19.3%, 15.7%, and 14.9% for the minimal, mild, moderate, and severely damaged images, respectively. For the blocked images, the PSNR increased by 161.9% after image restoration. In terms of MSE, the inpainting GAN model significantly reduced the values by 33.9%, 72.7%, 64.1%, and 61.5% for the minimal, mild, moderate, and severely damaged images, respectively. In the blocked images, we observed a 98% reduction in MSE. Finally, in the SSIM evaluation, the inpainting GAN model significantly improved the results with increases of 5.3%, 15.6%, 14.7%, and 14.8% for minimal, mild, moderate, and severely damaged images, respectively. For the blocked images, there was a 14.3% increase in the SSIM.

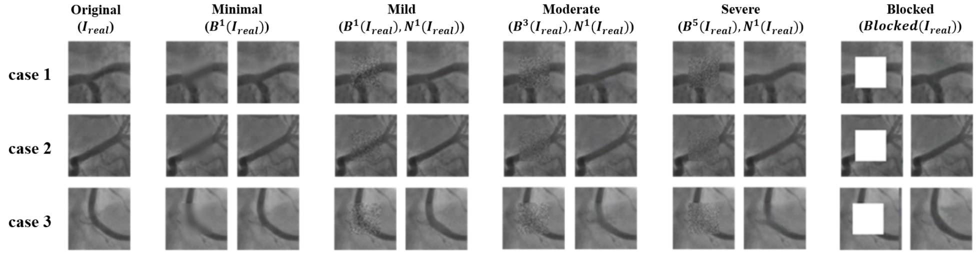

Figure 20. Examples of pre- and post-restoration using a inpainting GAN model. The image shows the results before and after restoration. To test the coronary restoration performance, samples with degradation levels ranging from minimal to completely blocked were selected.

The sample images in Fig. 20 demonstrate successful image restoration even in the case of blocked images. In Fig. 20 case 1, the bifurcation lesion was effectively restored even in severely damaged or blocked images. However, the degree of bifurcation in these images differed slightly from that of the original image. In Fig. 20 case 2, although the vessel diameter slightly decreased in the mildly damaged image compared with that in the original image, it was still well restored, even in the blocked image. Conversely, in Fig. 20 case 3, the vessel diameter in the blocked image exhibited a slight increase compared to that in the original image.

IV. DISCUSSION

The primary objective of this dissertation is to restore partially missing coronary arteries in both segmentation results and X-ray images. We have introduced two distinct methods to achieve this goal. The first method, detailed in Section II.1, was a post-processing to reconnect fragmented vessels in segmentation results using local geometric features. The second method, outlined in Section II.2, utilizes generative adversarial network (GAN) to reconstruct broken or noisy vessels in X-ray images. These methods are interrelated and serve the common purpose of restoring partially missing coronary arteries. In Section IV.1, we discuss into the reconnection methods, while in Section IV.2 we discuss the generative methods.

1. Discussion for Reconnection of Fragmented Parts of Coronary Arteries Using Local Geometric Features

In Section II.1, we proposed a robust method for reconnecting fragmented segments of coronary arteries in 2D X-ray angiography images, which are essential for coronary artery stenting procedures. Reconnecting fragmented blood vessels in a segmentation output can be invaluable in coronary artery procedures and diagnosis.

There are two main reasons why coronary arteries appear fragmented when obtained through segmentation of 2D X-ray angiography images. First, when the patient has a cardiovascular disease such as CTO in which the coronary artery is completely blocked, the lesion is not contrasted and is therefore invisible in the image. Second, the vesselness probability provided by a single NN model is not robust even

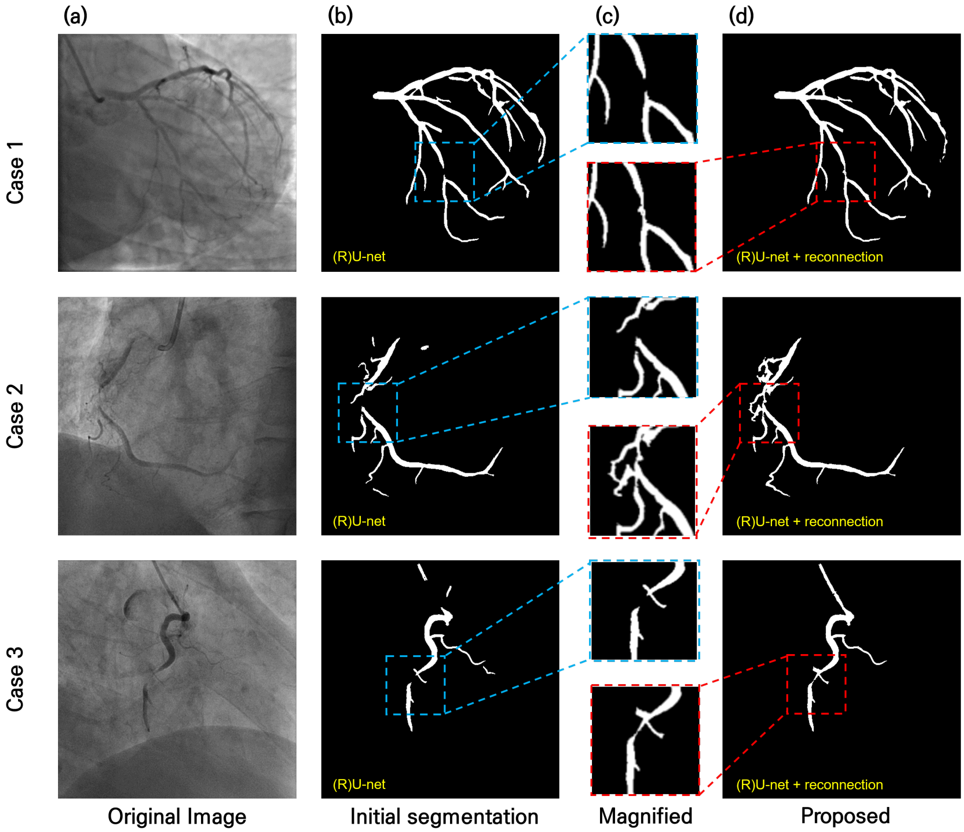

Figure 21. Reconnection results in 2D X-ray angiography of patients with interrupted vessel structures using the proposed method. (a) 2D X-ray angiography, (b) Initial segmentation results of interrupted vessel structure, (c) Magnified missing areas (blue) and reconnected areas (red), (d) Final reconnected segmentation results.

when the NN model is highly optimized. Consequently, no single threshold can exactly separate the foreground and background, especially when the target is elongated and thin, as is the case with coronary arteries.

The proposed method aims to reconnect disconnected fragments by combining both global and local CNN models. There can be cases where vesselness responses are weak when global CNN model is considered alone, even though the coronary artery actually exists. For this reason, fragmented regions can exist at multiple locations, and such cases occur frequently when targeting thin and elongated objects. The local CNN model can compensate for the missing regions by learning with local patch images that have balanced foreground and background pixel samples.

Specifically, the local patches can compensate for weak vesselness signals by considering the local properties of blood vessels. However, some challenges still exist. In the heart, there are many other tubular-shaped objects around the coronary arteries. Thus, sometimes the local patches may overcompensate around the coronary artery, which would increase false positives. Patch-based compensation is based on the premise that contrast values can be observed; however, it is sometimes impossible to observe the image values in cases of CTO like interrupted vessel structures. In this scenario, patch-based compensation is impossible.

To solve this problem, additional processing is required based on prior knowledge. Hence, the global and local CNN models are combined with a customized weighting function considering vascular dynamics. The vascular dynamics are described in terms of direction variations. The parameters of the SIG function were learned from a pool of centerlines manually annotated by medical experts. The purpose of the SIG function is to intuitively compensate for the missing regions in which contrast signals rarely exist in such interrupted vessel structures as shown in Fig. 21. The SIG function can be combined with two CNN models by providing artificial signals at the regions based on the expected central directions of the blood vessels.

A limitation of this method may arise when the length of the missing regions is too long. In such cases, the missing regions may not be included in the candidates for the reconnection process. In addition, if the radius value for the semicircle domain is too large, some unwanted false positives may be accompanied. Thus, it is necessary to set the radius parameter to an optimal value which is analyzed and suggested in our experiment. The local vascular dynamics in the semicircle region are linearly expressed in this investigation, which is difficult to properly represent some complex vascular shapes. Fortunately, the feasibility was demonstrated that the missing regions generated from the initial prediction obtained from the architectures of U-Net series were sufficiently recovered with the proposed method.

In the case that the contrast value at the missing region has only little difference from the background value due to the CTO lesion, the missing region may be overcompensated and some artifacts may accompany. However, even considering these side-effects, it can provide much more useful information to the clinical operator than when the region is completely missing. In addition, it is more advantageous when applying to the registration method between centerlines from 2D X-ray and 3D CT with the completely connected vessels without missing regions.

The fragmented arteries would likely be more accurately reconnected with higher-order curves rather than with a linear model. Our next challenge is to estimate the most plausible higher-order curve that represents the local centerline of the fragmented region. In general, the computational cost of estimating higher-order parameters may be very high. We are exploring methods to quickly estimate the parameters using a data-driven approach. In addition, the proposed method will be applied to the development of image-based procedure guidance and diagnosis systems.

2. Discussion for Reconstruction of Partially Broken Vascular Structure

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 interrupted vessel structure, 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.

However, as shown in Fig. 22, in cases such as 1, 2, and 3, we observed one limitation: vessels that were originally separated are merged into a single vessel in the reconstruction results. Additionally, in cases 4 and 5, there were other limitations, such as the creation of new pathways that did not exist in the original images. The possible causes of such distortions could be as follows. 1) During the training of the inpainting GAN network, at a certain point in the learning process, areas outside the mask become identical to the original image, and changes occur only in the masked areas. These areas are essential information for the discriminator, which determines whether the input image is real or generated. To robustly generate vessels that resemble real ones, it is necessary to train a local discriminator that distinguishes real images from generated images based only on the mask area. However, in the proposed method, only a global discriminator that discriminates the entire 128×128 image was trained, and a local discriminator for the 64×64 mask area was not included.

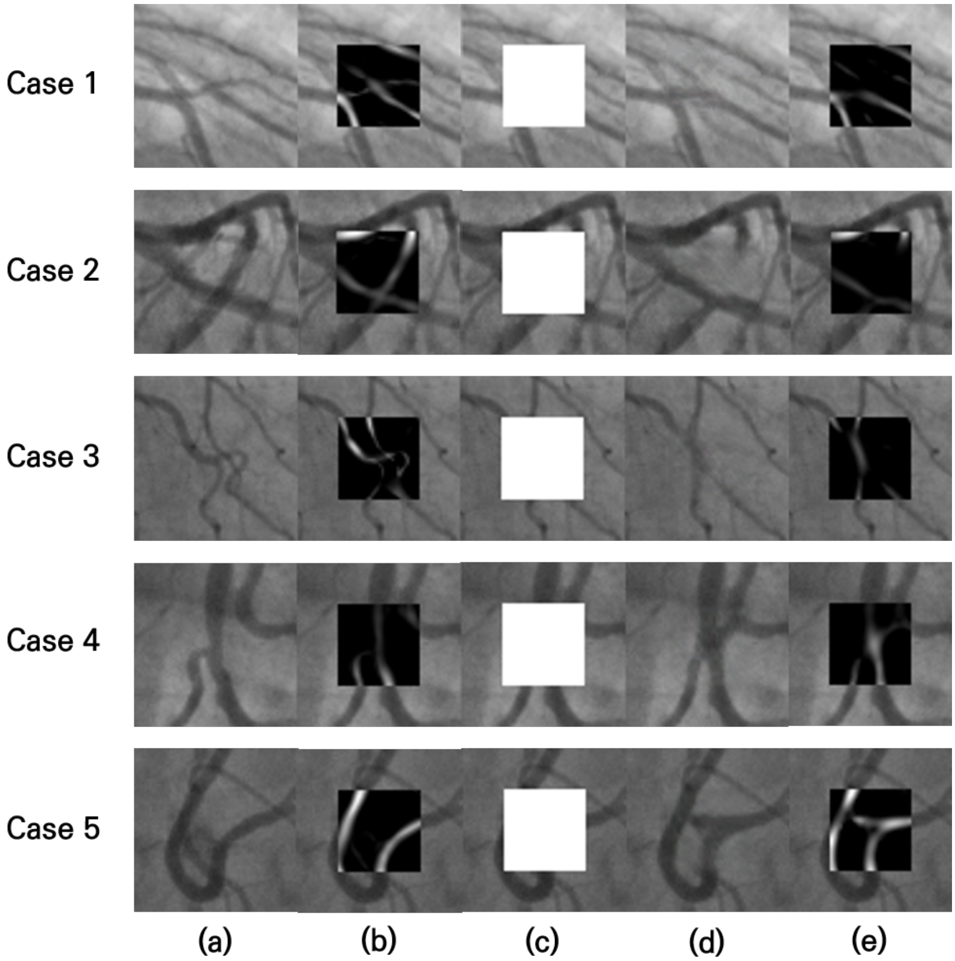

Figure 22. Limitations of the proposed reconstruction methods for specific cases. (a) original X-ray image patches, (b) vesselness map of original X-ray patches, (c) model input patches (d) reconstructed patches (e) vesselness map of reconstructed patches.

Therefore, this distortion could be observed due to the absence of the local discriminator. 2) The model might have lacked the capacity to generate coronary arteries with complex shapes. 3) Finally, the hyper-parameters used for training may not have been sufficiently fine-tuned. To generate even more realistic vascular images, each possible cause will be more intricately analyzed in future researches.

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.

The proposed method, meanwhile, 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, 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 interrupted vessel structure cases. We expect the proposed method to be utilized for image-guided procedure or diagnosis systems in real clinical sites.

3. Discussion for Demonstrating Clinical Feasibility

In this pilot study, we demonstrate the feasibility of using an inpainting GAN to restore images with varying degrees of damage or even deleted segments in ICA. Our deep learning-based model enabled instant restoration and improvement of ICA image quality. This approach holds the potential for application in coronary intervention procedures, especially in emergency settings or interventions involving interrupted vessel structures.

X-ray coronary angiography serves as a confirmatory test for CAD and provides essential guidance for PCI procedures. However, the inherent limitation of two-dimensional projection images make it challenging to accurately assess the severity of lesions, particularly in cases of overlapping tortuous vessels. Additionally, the inability to visualize vessels from different angles can hinder accurate assessments. Moreover, in the distal part of a CTO, vessels are not visible because the contrast agent cannot reach them. Registration-based methods have been proposed to estimate the original structure of lost vessels in X-ray images. These methods aim to match the 2D and 3D coronary centerlines extracted from X-ray and CT images. However, these techniques have limitations. One is the need to extract new centerlines when the X-ray angle changes. Additionally, acquiring supplementary 3D CT images beyond the original X-ray images is challenging. In contrast, our model does not require additional images from CT or other modalities to address these limitations.

In contrast, our proposed method relies exclusively on the 2D X-ray modality and can restore coronary arteries with varying degrees of degradation or even complete absence, which are commonly encountered in real-world operational scenarios. Consequently, it can offer guidance for interventional procedures while simultaneously estimating damaged or missing vessel segments.

This study had several limitations. First, in some cases, the restoration process results in slight distortions of the vascular pathway or vessel diameter. However, preservation of the vessel centerline, which is critical for guiding interventional procedures, appears to be adequate. Second, in this pilot study, we evaluated artificially degraded or blocked images rather than actual patient vessels with more complex diseases or interrupted vessel structures. Therefore, the effectiveness of the proposed model on authentic images remains uncertain. Further studies incorporating a diverse range of diseased vessels and interrupted vessel structures are necessary to investigate the applicability of these findings.

In conclusion, we demonstrated that inpainting GAN restores ICA images with varying damage levels, including deletions. Our deep learning-based model enables immediate restoration and enhancement of ICA image quality. This approach has the potential for application in coronary intervention procedures, particularly in emergency settings or PCI.

V. CONCLUSION

2D X-ray angiography images serve as the gold standard image modality for performing percutaneous coronary intervention procedures on patients with coronary artery disease. However, challenges arise in accurately identifying coronary artery regions within 2D X-ray angiography images due to issues such as noise and cardiac motion artifacts. Furthermore, in cases of patients with rapid heart rates or obesity, obtaining clear vascular imaging can be difficult, posing obstacles to delineating vascular pathways within the images. Additionally, in scenarios of chronic total occlusion where a blood vessel is completely blocked, the absence of contrast agent passage results in the invisible depiction of coronary artery regions in the 2D X-ray angiography images. This absence makes referencing coronary artery areas during percutaneous coronary intervention procedures considerably more challenging.

There are numerous ways that interventional operators can use to help them during procedures: 1) Accurately segmenting coronary artery regions within 2D X-ray angiography images, or registering segmentation results with 3D CT images and presenting them to the operator, allowing easy identification of coronary artery regions. 2) Reconstructing noised or invisible vessels in 2D X-ray angiography images, allowing operators to timely see coronary artery areas straight from the 2D X-ray images.

However, despite using state-of-the-art deep-learning-based segmentation models to segment coronary arteries, the limitations of single CNN models have led to partially fragmented results when segmenting complex coronary artery regions. Additionally, in cases such as chronic total occlusion (CTO), where contrast agents cannot pass due to lesions, resulting in a lack of vessel information in the image, partially broken segmentation results have also been obtained. The issue of partially broken segmentation results has rarely been discussed and studied.

In this dissertation, as the first approach to support interventional procedures, a sophisticated post-processing method utilizing local geometric features is proposed to reconnect fragmented coronary artery segmentation results. In this dissertation, as the first approach to support interventional procedures, a sophisticated post-processing method utilizing local geometric features is proposed to reconnect fragmented coronary artery segmentation results. Through the proposed method, it was observed that the partially broken coronary artery segmentation results, due to the limitations of a single CNN model as well as interrupted vessel structure, were smoothly reconnected along the vessel's direction. Furthermore, significant quantitative and qualitative improvements were observed across all approaches when the suggested method was applied to the initial segmentation results of state-of-the-art methods.

Furthermore, as the second approach, a GAN based novel Multi-Attention Block (MAB) network was proposed for the reconstruction of vascular areas solely from single 2D X-ray angiography images. Moreover, a vesselness-loss function is introduced to optimally reconstruct objects with structures such as thin and elongated vessels. To assess the performance of the proposed method across various cases that may encounter in real clinical settings, original images were artificially degraded and then reconstructed. The results demonstrated that the proposed methods realistically restored the centerline of blood vessels to a level suitable for reference during image-guided procedures.

The methods proposed in this dissertation, including the segmentation-based reconnection methods and the generative model-based reconstruction method, are verified to be capable of sufficiently restoring the missing coronary arteries to a level that can be referenced by operators during interventional procedures. The proposed methods are expected to be used in real clinical settings in the near future.

REFERENCES

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

[2] Sanchis-Gomar F., Perez-Quilis C., Leischik R., and Lucia A., "Epidemiology of coronary heart disease and acute coronary syndrome," *Annals of translational medicine*, vol. 4, no. 13, 2016.

[3] Mensah G. A., Roth G. A., and Fuster V., *The global burden of cardiovascular diseases and risk factors: 2020 and beyond*, *Journal of the American College of Cardiology*, vol. 74, 20, pp. 2529–2532, 2019.

[4] Stary H. C., Chandler A. B., Dinsmore R. E., Fuster V., Glagov S., Insull Jr W., *et al.*, "A definition of advanced types of atherosclerotic lesions and a histological classification of atherosclerosis: A report from the committee on vascular lesions of the council on arteriosclerosis, american heart association," *Circulation*, vol. 92, no. 5, pp. 1355–1374, 1995.

[5] Baka N., Metz C., Schultz C., Neefjes L., Geuns R. J. van, Lelieveldt B. P., *et al.*, "Statistical coronary motion models for 2d+ t/3d registration of x-ray coronary angiography and cta," *Medical image analysis*, vol. 17, no. 6, pp. 698–709, 2013.

[6] Baka N., Lelieveldt B., Schultz C., Niessen W., and Walsum T. van, "Respiratory motion estimation in x-ray angiography for improved guidance dur-

ing coronary interventions," *Physics in Medicine & Biology*, vol. 60, no. 9, p. 3617, 2015.

[7] Zhu J., Li H., Ai D., Yang Q., Fan J., Huang Y., *et al.*, "Iterative closest graph matching for non-rigid 3d/2d coronary arteries registration," *Computer Methods and Programs in Biomedicine*, vol. 199, p. 105 901, 2021.

[8] Park T., Khang S., Jeong H., Koo K., Lee J., Shin J., *et al.*, "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, 2022.

[9] Jeong D., Kim D., Ryu J., and Cho K. H., "Deep-learning-based registration of diagnostic angiogram and live fluoroscopy for percutaneous coronary intervention," *IEEE Access*, vol. 9, pp. 103 465–103 480, 2021.

[10] Wu W., Zhang J., Peng W., Xie H., Zhang S., and Gu L., "Car-net: A deep learning-based deformation model for 3d/2d coronary artery registration," *IEEE Transactions on Medical Imaging*, 2022.

[11] Kalisz K., Buethe J., Saboo S. S., Abbara S., Halliburton S., and Rajiah P., "Artifacts at cardiac ct: Physics and solutions," *Radiographics*, vol. 36, no. 7, pp. 2064–2083, 2016.

[12] Frangi A. F., Niessen W. J., Vincken K. L., and Viergever M. A., "Multiscale vessel enhancement filtering," in *International conference on medical image computing and computer-assisted intervention*, Springer, 1998, pp. 130–137.

[13] Krissian K., Malandain G., Ayache N., Vaillant R., and Trousset Y., "Model-based detection of tubular structures in 3d images," *Computer vision and image understanding*, vol. 80, no. 2, pp. 130–171, 2000.

[14] Li Q., Sone S., and Doi K., “Selective enhancement filters for nodules, vessels, and airway walls in two-and three-dimensional ct scans,” *Medical physics*, vol. 30, no. 8, pp. 2040–2051, 2003.

[15] Lin Q., “Enhancement, detection, and visualization of 3d volume data,” Ph.D. dissertation, Linköping University, Department of Electrical Engineering, 2001.

[16] Lorenz C., Carlsen I.-C., Buzug T. M., Fassnacht C., and Weese J., “Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2d and 3d medical images,” in *CVRMed-MRCAS'97*, Springer, 1997, pp. 233–242.

[17] Yuan Y. and Chung A. C., “Multi-scale model-based vessel enhancement using local line integrals,” in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2008, pp. 2225–2228.

[18] Kass M., Witkin A., and Terzopoulos D., “Snakes: Active contour models,” *International journal of computer vision*, vol. 1, no. 4, pp. 321–331, 1988.

[19] Nirmala Devi S and Kumaravel N, “Comparison of active contour models for image segmentation in x-ray coronary angiogram images,” *Journal of medical engineering & technology*, vol. 32, no. 5, pp. 408–418, 2008.

[20] Taghizadeh Dehkordi M., Doost Hoseini A. M., Sadri S., and Soltanianzadeh H., “Local feature fitting active contour for segmenting vessels in angiograms,” *IET Computer Vision*, vol. 8, no. 3, pp. 161–170, 2014.

[21] Chen X., Williams B. M., Vallabhaneni S. R., Czanner G., Williams R., and Zheng Y., “Learning active contour models for medical image segmentation,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 11 632–11 640.

[22] Gu J., Fang Z., Gao Y., and Tian F., "Segmentation of coronary arteries images using global feature embedded network with active contour loss," *Computerized Medical Imaging and Graphics*, vol. 86, p. 101 799, 2020.

[23] Ronneberger O., Fischer P., and Brox T., "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, pp. 234–241.

[24] Siddique N., Sidike P., Elkin C., and Devabhaktuni V., "U-net and its variants for medical image segmentation: Theory and applications," *arXiv preprint arXiv:2011.01118*, 2020.

[25] Yang S., Kweon J., Roh J.-H., Lee J.-H., Kang H., Park L.-J., *et al.*, "Deep learning segmentation of major vessels in x-ray coronary angiography," *Scientific reports*, vol. 9, no. 1, pp. 1–11, 2019.

[26] Wu W., Zhang J., Xie H., Zhao Y., Zhang S., and Gu L., "Automatic detection of coronary artery stenosis by convolutional neural network with temporal constraint," *Computers in biology and medicine*, vol. 118, p. 103 657, 2020.

[27] Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., *et al.*, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.

[28] Karras T., Laine S., and Aila T., "A style-based generator architecture for generative adversarial networks," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 4401–4410.

[29] Zhu J.-Y., Park T., Isola P., and Efros A. A., "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017.

[30] Pathak D., Krahenbuhl P., Donahue J., Darrell T., and Efros A. A., "Context encoders: Feature learning by inpainting," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2536–2544.

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

[32] Zeng Y., Fu J., Chao H., and Guo B., "Aggregated contextual transformations for high-resolution image inpainting," *IEEE Transactions on Visualization and Computer Graphics*, 2022.

[33] Armanious K., Kumar V., Abdulatif S., Hepp T., Gatidis S., and Yang B., "Ipa-medgan: Inpainting of arbitrary regions in medical imaging," in *2020 IEEE international conference on image processing (ICIP)*, IEEE, 2020, pp. 3005–3009.

[34] Sim B., Oh G., Kim J., Jung C., and Ye J. C., "Optimal transport driven cycle-gan for unsupervised learning in inverse problems," *SIAM Journal on Imaging Sciences*, vol. 13, no. 4, pp. 2281–2306, 2020.

[35] Wang Q., Chen Y., Zhang N., and Gu Y., "Medical image inpainting with edge and structure priors," *Measurement*, vol. 185, p. 110 027, 2021.

[36] Bertalmio M., Sapiro G., Caselles V., and Ballester C., "Image inpainting," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, 2000, pp. 417–424.

[37] Efros A. A. and Freeman W. T., "Image quilting for texture synthesis and transfer," in *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, 2001, pp. 341–346.

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

[39] Hays J. and Efros A. A., "Scene completion using millions of photographs," *ACM Transactions on Graphics (ToG)*, vol. 26, no. 3, 4–es, 2007.

[40] Darabi S., Shechtman E., Barnes C., Goldman D. B., and Sen P., "Image melding: Combining inconsistent images using patch-based synthesis," *ACM Transactions on graphics (TOG)*, vol. 31, no. 4, pp. 1–10, 2012.

[41] Iizuka S., Simo-Serra E., and Ishikawa H., "Globally and locally consistent image completion," *ACM Transactions on Graphics (ToG)*, vol. 36, no. 4, pp. 1–14, 2017.

[42] Yu J., Lin Z., Yang J., Shen X., Lu X., and Huang T. S., "Generative image inpainting with contextual attention," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 5505–5514.

[43] Liu H., Jiang B., Xiao Y., and Yang C., "Coherent semantic attention for image inpainting," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 4170–4179.

[44] Yu J., Lin Z., Yang J., Shen X., Lu X., and Huang T. S., "Free-form image inpainting with gated convolution," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 4471–4480.

[45] Ma Y., Liu X., Bai S., Wang L., He D., and Liu A., "Coarse-to-fine image inpainting via region-wise convolutions and non-local correlation.," in *IJCAI*, 2019, pp. 3123–3129.

[46] Yi Z., Tang Q., Azizi S., Jang D., and Xu Z., "Contextual residual aggregation for ultra high-resolution image inpainting," in *Proceedings of the IEEE/CVF*

Conference on Computer Vision and Pattern Recognition, 2020, pp. 7508–7517.

[47] Guo X., Yang H., and Huang D., "Image inpainting via conditional texture and structure dual generation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021, pp. 14 134–14 143.

[48] Feng Z., Chi S., Yin J., Zhao D., and Liu X., "A variational approach to medical image inpainting based on mumford-shah model," in *2007 International Conference on Service Systems and Service Management*, IEEE, 2007, pp. 1–5.

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

[50] Arnold M., Ghosh A., Ameling S., and Lacey G., "Automatic segmentation and inpainting of specular highlights for endoscopic imaging," *EURASIP Journal on Image and Video Processing*, vol. 2010, pp. 1–12, 2010.

[51] Armanious K., Mecky Y., Gatidis S., and Yang B., "Adversarial inpainting of medical image modalities," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2019, pp. 3267–3271.

[52] Han K., Jeon J., Jang Y., Jung S., Kim S., Shim H., *et al.*, "Reconnection of fragmented parts of coronary arteries using local geometric features in x-ray angiography images," *Computers in biology and medicine*, vol. 141, p. 105 099, 2022.

[53] Fan J., Yang J., Wang Y., Yang S., Ai D., Huang Y., *et al.*, "Multichannel fully convolutional network for coronary artery segmentation in x-ray angiograms," *Ieee Access*, vol. 6, pp. 44 635–44 643, 2018.

[54] Jo K., Kweon J., Kim Y.-H., and Choi J., "Segmentation of the main vessel of the left anterior descending artery using selective feature mapping in coronary angiography," *IEEE Access*, vol. 7, pp. 919–930, 2018.

[55] Xia S., Zhu H., Liu X., Gong M., Huang X., Xu L., *et al.*, "Vessel segmentation of x-ray coronary angiographic image sequence," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 5, pp. 1338–1348, 2019.

[56] Weng Y., Zhou T., Li Y., and Qiu X., "Nas-unet: Neural architecture search for medical image segmentation," *IEEE Access*, vol. 7, pp. 44 247–44 257, 2019.

[57] Zhao X., Zhang P., Song F., Fan G., Sun Y., Wang Y., *et al.*, "D2a u-net: Automatic segmentation of covid-19 lesions from ct slices with dilated convolution and dual attention mechanism," *arXiv preprint arXiv:2102.05210*, 2021.

[58] Jégou S., Drozdzal M., Vazquez D., Romero A., and Bengio Y., "The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2017, pp. 11–19.

[59] Dodge Jr J. T., Brown B. G., Bolson E. L., and Dodge H. T., "Lumen diameter of normal human coronary arteries. influence of age, sex, anatomic variation, and left ventricular hypertrophy or dilation.," *Circulation*, vol. 86, no. 1, pp. 232–246, 1992.

[60] Lesage D., Angelini E. D., Bloch I., and Funka-Lea G., "Medial-based bayesian tracking for vascular segmentation: Application to coronary arteries in 3d

ct angiography," in *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, IEEE, 2008, pp. 268–271.

[61] Han K., Koo H., Jung S., Park H.-B., Hong Y., Shim H., *et al.*, "Reconstruction of partially broken vascular structures in x-ray images via vesselness-loss-based multi-scale generative adversarial networks," *IEEE Access*, 2023.

[62] Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[63] Xie S., Girshick R., Dollár P., Tu Z., and He K., "Aggregated residual transformations for deep neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1492–1500.

[64] Liu Z., Mao H., Wu C.-Y., Feichtenhofer C., Darrell T., and Xie S., "A convnet for the 2020s," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 11 976–11 986.

[65] He K., Zhang X., Ren S., and Sun J., "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[66] Veit A., Wilber M. J., and Belongie S., "Residual networks behave like ensembles of relatively shallow networks," *Advances in neural information processing systems*, vol. 29, 2016.

[67] Isola P., Zhu J.-Y., Zhou T., and Efros A. A., "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.

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

[69] Lim J. H. and Ye J. C., "Geometric gan," *arXiv preprint arXiv:1705.02894*, 2017.

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

[71] Ledig C., Theis L., Huszár F., Caballero J., Cunningham A., Acosta A., *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681–4690.

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

[73] Sato Y., Nakajima S., Atsumi H., Koller T., Gerig G., Yoshida S., *et al.*, "3d multi-scale line filter for segmentation and visualization of curvilinear structures in medical images," in *CVRMed-MRCAS'97*, Springer, 1997, pp. 213–222.

[74] Yang J., Huang M., Fu J., Lou C., and Feng C., "Frangi based multi-scale level sets for retinal vascular segmentation," *Computer Methods and Programs in Biomedicine*, vol. 197, p. 105 752, 2020.

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

[76] Lee H., Shim H., and Chang H.-J., "Intensity-vesselness gaussian mixture model (ivgmm) for 2d+ t segmentation of coronary arteries for x-ray angiography image sequences," *Journal of X-ray Science and Technology*, vol. 23, no. 5, pp. 579–592, 2015.

[77] Abdullah-Al-Mamun M., Tyagi V., and Zhao H., "A new full-reference image quality metric for motion blur profile characterization," *IEEE Access*, vol. 9, pp. 156 361–156 371, 2021.

ABSTRACT (IN KOREAN)

2D X-선 혈관 조영술 영상에서의

딥러닝기반 유실된 혈관 복원 기법

< 지도교수 장혁재 >

연세대학교 대학원 의과학과

한경훈

목적 - 관상동맥 시술은 주로 X-선 혈관조영 영상을 기반으로 시행된다. 그러나, X-선 혈관조영 영상 내의 노이즈 또는 심장 모션 아티팩트는 영상 내 관상동맥을 정확히 파악하는 것을 어렵게 만들기 때문에 영상기반 카테터 시술 시 많은 어려움이 존재한다. 뿐만 아니라, 관상동맥이 완전히 폐색된 만성 완전 폐색 (CTO)의 경우 조영제가 통과되지 않아 혈관의 주행 경로를 알지 못하여 무리한 와이어 조작으로 인해 빈번하게 혈관 천공 및 응급 상황이 발생할 수 있는 문제점이 존재한다. 이러한 문제점을 해결하기 위하여 영상기반 가이딩 시술을 보조하기 위한 딥러닝 기반 프레임워크를 제안한다.

관상동맥 시술은 주로 2차원 (2D) X-선 혈관 조영 영상을 기반으로 시행된다. 그러나 X-선 혈관 조영 영상 내의 노이즈나 심장 모션 아티팩트는 영상에

서 관상동맥을 정확하게 식별하는 것을 어렵게 만들기 때문에, 영상 기반 카테터 시술에서 많은 어려움이 존재한다. 뿐만 아니라, 관상동맥이 완전히 폐쇄된 만성 완전 폐색 (CTO)의 경우 조영제가 통과되지 않아 혈관의 경로를 알 수 없으며, 과도한 와이어 조작으로 인해 혈관 천공 및 응급 상황이 빈번히 발생할 수 있는 문제점이 존재한다. 본 논문에서는, 영상 유도 절차를 지원하기 위해 파손된 관상 동맥 분할 결과를 다시 연결하거나 2차원 X-선 혈관 조영 이미지를 재구성하여 누락된 관상 동맥을 복원하기 위한 딥 러닝 기반 프레임워크를 제안한다.

방법 - 영상 내 관상동맥을 강건하게 분할할 수 있는 모델은 시술자에게 보다 정확한 관상동맥 영역 및 혈관 경로를 제공할 수 있으므로 임상적으로 유용하다. 해당 모델은 영상과 영상 내 관상동맥에 해당하는 참 값의 쌍을 통한 지도 학습을 통해 얻어진다. 하지만 단일 합성곱 신경망 (CNN) 모델은 관상동맥의 복잡하고 얇은 구조를 분할하는데 한계가 존재하여, 종종 관상동맥 분할 시 부분적으로 끊어진 결과를 출력하게 된다. 뿐만 아니라 만성 완전 폐색 등의 병변으로 인하여 영상 내 혈관 정보가 존재하지 않는 문제로 인하여 하나로 연결되어 있어야하는 관상동맥이 부분적으로 유실되어 분할되게 된다. 본 연구에서는 지역 정보 및 기하학적 사전지식을 활용하여 부분적으로 끊어진 관상동맥 영역을 재연결하기 위한 정교한 후처리 기법을 제안하였다.

뿐만 아니라, 영상 내 잡음 또는 병변으로 인하여 혈관 조영이 선명하게 이루

어지지 않아 혈관 주행 경로가 명확히 구분되지 않는 영상에서 혈관 영역을 선명하게 복원할 수 있는 모델은 시술자에게 보다 정확한 혈관 경로를 제공할 수 있으므로 임상적으로 유용하다. 해당 모델은 X-선 영상만을 사용하는 비지도 학습을 통해 얻어질 수 있다. 본 논문에서는 어떠한 다른 모달리티에도 의존하지 않고 오직 X-선 영상만을 기반으로 혈관 영역을 복원하는 생성적 적대 신경망 (GAN) 모델 기반의 혈관 복원 기법을 제안하였다.

결과 - 먼저, 끊어진 관상동맥 분할결과를 재연결하기 위해 제안된 후처리 기법의 성능을 테스트 해본 결과, 초기 분혈 결과에서 영상 당 2.308개 존재하였던 끊어진 영역이 제안 기법을 통해 1.197개로 크게 줄어드는 것을 확인하였다. 뿐만 아니라, 최신의 분할 모델의 초기 분할 결과에 제안된 후처리 기법을 적용한 결과 지역적으로 끊어진 영역에서 혈관 체적 일치도가 평균 약 2.33배 증가하였으며, 자카드 인덱스의 경우 평균 약 2.88배 증가하는 것을 확인하였다. 또한, 제안된 혈관 복원 결과의 경우 다른 최신의 복원 모델들에 비하여 3개의 영상 품질 측정 지표에서 가장 우수한 성능을 보이는 것을 정량적으로 확인하였으며, 가이딩 중재 시술시 중요한 혈관의 중심선을 잘 보존하여 혈관을 복원한 것을 정성적으로 확인하였다.

결론 - 본 학위논문에서는 관상동맥 중재 시술 시 시술자가 참조할 수 있는 영상을 제공하기 위한 딥러닝 기반의 기법들을 제안하였다. 이를 위하여, 혈

관 분할 결과 시 끊어진 영역을 재연결하기 위해 지역적 정보를 활용한 후처리 기법과, 영상 내 혈관 주행 경로가 명확하지 않을 때 혈관 영역을 복원하기 위한 영상 생성 기법을 제안하였다. 제안된 각각의 기법은 시술자가 시술시 참조할 수 있을만한 수준의 혈관 분할 결과 또는 X-선 영상을 제공한다는 점에서, 우리는 제안된 기법이 가까운 시일 내 실제 임상 현장에 활용될 수 있을 것이라 기대된다.

핵심되는 말: 딥러닝, 2D 혈관조영영상, 관상동맥, 만성 완전 폐색, 심혈관 질환